

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 07:52:30 ; Search time 1498 Seconds  
(without alignments)  
518.880 Million cell updates/sec

Title: US-09-606-804-1  
Perfect score: 19  
Sequence: 1 ccaatnnnnnnnnccacg 19

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl :

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rtd.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	10	52.6	19	6	AR274677	Sequence AR274677
2	10	52.6	19	6	AX275341	Sequence AX275341
3	10	52.6	19	6	BD016693	Method fo BD016693
4	10	52.6	19	6	E28797	Endoplasmic E28797
5	10	52.6	19	6	E28801	Endoplasmic E28801
6	10	52.6	19	6	E28802	Endoplasmic E28802
7	10	52.6	19	6	E28803	Endoplasmic E28803
8	10	52.6	19	6	E28804	Endoplasmic E28804
9	10	52.6	19	6	E28805	Endoplasmic E28805
10	10	52.6	19	6	E28806	Endoplasmic E28806
11	10	52.6	19	6	E28807	Endoplasmic E28807
12	10	52.6	19	6	E28808	Endoplasmic E28808
13	10	52.6	19	6	E28809	Endoplasmic E28809
14	10	52.6	19	6	E44202	Endoplasmic E44202
15	10	52.6	19	6	E44206	Endoplasmic E44206
16	10	52.6	19	6	E44207	Endoplasmic E44207
17	10	52.6	19	6	E44208	Endoplasmic E44208
18	10	52.6	19	6	E44209	Endoplasmic E44209
19	10	52.6	19	6	E44210	Endoplasmic E44210
20	10	52.6	19	6	E44211	Endoplasmic E44211
21	10	52.6	19	6	E44212	Endoplasmic E44212
22	10	52.6	19	6	E44213	Endoplasmic E44213
23	10	52.6	19	6	E44214	Endoplasmic E44214
24	10	52.6	24	6	AX443752	Sequence AX443752
25	10	52.6	25	6	AX447733	Sequence AX447733
26	10	52.6	29	6	E44234	Endoplasmic E44234
27	10	52.6	40	6	BD180759	Array of BD180759
28	10	52.6	99	9	HUMCG3A02	Human alpha M10797
29	10	52.6	105	8	AF479941	Saccharom AF479941
30	10	52.6	112	6	AX343726	Sequence AX343726
31	10	52.6	118	8	ATH524636	Arabidops ATH524636
32	10	52.6	122	6	E28800	Endoplasmic E28800
33	10	52.6	122	6	E44205	Endoplasmic E44205
34	10	52.6	125	14	AF271990	Human ade AF271990
35	10	52.6	127	6	AX343728	Sequence AX343728
36	10	52.6	133	8	ATH523697	Arabidops ATH523697
37	10	52.6	133	8	ATH523876	Arabidops ATH523876
38	10	52.6	136	6	AX033516	Sequence AX033516
39	10	52.6	136	6	AX033517	Sequence AX033517
40	10	52.6	138	8	AY203001	Arabidops AY203001
41	10	52.6	142	6	AX033504	Sequence AX033504
42	10	52.6	143	6	AX033505	Sequence AX033505
43	10	52.6	143	9	D10339505	Sequence D10339505
44	10	52.6	145	6	AX072778	Sequence AX072778
45	10	52.6	149	6	AX033508	Sequence AX033508

ALIGNMENTS

RESULT 1	AR274677	Sequence 1 from patent US 6506600.	19 bp	DNA	linear	PAT 10-APR-2003
LOCUS	AR274677					
DEFINITION	Sequence 1 from patent US 6506600.					
ACCESSION	AR274677					
VERSION	AR274677.1	GI:29707222				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					

REFERENCE  
1 (bases 1 to 19)  
Heronat, P.L., Mane, M. and Liu, Y.  
Secreting products from skin by adeno-associated virus (AAV) gene transfer  
Patent: US 6506600-A 1 14-JAN-2003;

```
FEATURES
  source
    Location/Qualifiers
      1..19
      /organism="unknown"
BASE COUNT      3 a      5 c      1 g      1 t      9 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 10; DB 6; Length 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 1 CCAATNNNNNNNNCCACG 19

RESULT 2
AX275341
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
  source
    Location/Qualifiers
      1..19
      /organism="unidentified"
      /mol_type="genomic DNA"
      /db_xref="taxon:32644"
      /note="glucose response element"
BASE COUNT      3 a      5 c      1 g      1 t      9 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 10; DB 6; Length 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 1 CCAATNNNNNNNNCCACG 19

RESULT 3
BD016693
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
  OS Artificial Sequence
  PN JP 2001238699-A/2
  PD 04-SEP-2001
  PF 01-MAR-2000 JP 2000055384
  PI RIEKO AKAZAWA,SUMIKO KANEDA,HIDEKI YANAGI,TAKASHI YURA PC
    C12Q1/68,A61K45/00,A61P1/04,A61P3/10,A61P9/10,A61P17/02,A61P25/ PC
    28,
  PC A61P35/00,A61P37/00,C12N5/10,C12N15/09//((C12N5/10,C12R1:91),
    (C12N15/09,C12R1:91),C12N5/00,C12N15/00,C12N5/00,C12R1:91),
    (C12N15/00,C12R1:91))
```

```
PC (C12N15/00,C12R1:91)
CC Description of Artificial Sequence:The sequence as shown in
CC SEQ ID NO: 2
CC is ERSE concensus sequence.
CC n is A or C or G or T.
FH Key
  Location/Qualifiers
    1..19
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
BASE COUNT      3 a      5 c      1 g      1 t      9 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 10; DB 6; Length 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 1 CCAATNNNNNNNNCCACG 19

RESULT 4
E28797
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
  OS Unidentified
  PN JP 1999243959-A/1
  PD 14-SEP-1999
  PF 04-MAR-1998 JP 1998052453
  PR
  PI HIDEO YOSHIDA,HIDEKI YANAGI,TAKASHI YURA
  PC C12N15/09,A61K35/74,A61K35/76,A61K38/00,A61K48/00,C12N15/00,
  PC A61K37/02
  CC Strandedness: Double;
  CC Topology: Linear;
  FH Key
  FT enhancer
    Location/Qualifiers
      1..19
      /organism="unidentified"
      /mol_type="genomic DNA"
      /db_xref="taxon:32644"
BASE COUNT      3 a      5 c      1 g      1 t      9 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 10; DB 6; Length 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 1 CCAATNNNNNNNNCCACG 19

RESULT 5
E28801
LOCUS
DEFINITION
ACCESSION
VERSION
  E28801.1 GI:13020855
```

```

KEYWORDS      JP 1999243959-A/5.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens

REFERENCE
AUTHORS      Hideo Yoshida, Hideki Yanagi, Takashi Yura
TITLE        C12N15/09, A61K35/74, A61K35/76, A61K38/00, A61K48/00, C12N15/00,
JOURNAL      A61K37/02
COMMENT      Endoplasmic reticulum stress-response regulatory element
              Patent: JP 1999243959-A 5 14-SEP-1999;
              HSP RESEARCH INST INC

OS           Homo sapiens (human)
PN           JP 1999243959-A/5
PD           14-SEP-1999
PF           04-MAR-1998 JP 1998052453
PR           HIDEO YOSHIDA, HIDEKI YANAGI, TAKASHI YURA
PI           C12N15/09, A61K35/74, A61K35/76, A61K38/00, A61K48/00, C12N15/00,
PC           A61K37/02
CC           Strandedness: Double;
CC           Topology: Linear;
CC           Key      Location/Qualifiers
FT           enhancer 1..19.
              Location/Qualifiers
              1..19
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              2 t

BASE COUNT   3 a 9 c 5 g
ORIGIN

Query Match  52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 52.6%; Pred. No. 2.7e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

YQ 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 1 CCAATCGGGGCTCCACG 19
    |||||

RESULT 7
E28803
LOCUS      E28803
DEFINITION Endoplasmic reticulum stress-response regulatory element.
ACCESSION E28803
VERSION   E28803.1 GI:13020857
KEYWORDS  JP 1999243959-A/7.
SOURCE    Rattus sp.
ORGANISM  Rattus sp.
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
          1 (bases 1 to 19)
          Hideo, Y., Hideki, Y. and Takashi, Y.
          Endoplasmic reticulum stress-response regulatory element
          Patent: JP 1999243959-A 7 14-SEP-1999;
          HSP RESEARCH INST INC
          OS Rattus sp. (rat)
          PN JP 1999243959-A/7
          PD 14-SEP-1999
          PF 04-MAR-1998 JP 1998052453
          PR HIDEO YOSHIDA, HIDEKI YANAGI, TAKASHI YURA
          PC C12N15/09, A61K35/74, A61K35/76, A61K38/00, A61K48/00, C12N15/00,
          PC A61K37/02
          CC Strandedness: Double;
          CC Topology: Linear;
          CC Key      Location/Qualifiers
          FH enhancer 1..19.
          FT enhancer 1..19.
              Location/Qualifiers
              1..19
              /organism="Rattus sp."
              /mol_type="genomic DNA"
              /db_xref="taxon:10118"
              2 t

BASE COUNT   4 a 8 c 5 g
ORIGIN

Query Match  52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 52.6%; Pred. No. 2.7e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

YQ 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 1 CCAATCGGGGCTCCACG 19
    |||||

RESULT 8
E28804
LOCUS      E28804
DEFINITION Endoplasmic reticulum stress-response regulatory element.
ACCESSION E28804
VERSION   E28804.1 GI:13020858
KEYWORDS  JP 1999243959-A/8.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 19)
          Hideo, Y., Hideki, Y. and Takashi, Y.
          Endoplasmic reticulum stress-response regulatory element
          Patent: JP 1999243959-A 8 14-SEP-1999;
          HSP RESEARCH INST INC

OS           Homo sapiens (human)
PN           JP 1999243959-A/8
PD           14-SEP-1999
PF           04-MAR-1998 JP 1998052453
PR           HIDEO YOSHIDA, HIDEKI YANAGI, TAKASHI YURA
PI           C12N15/09, A61K35/74, A61K35/76, A61K38/00, A61K48/00, C12N15/00,
PC           A61K37/02
CC           Strandedness: Double;
CC           Topology: Linear;
CC           Key      Location/Qualifiers
FT           enhancer 1..19.
              Location/Qualifiers
              1..19
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:10095"
              2 t

BASE COUNT   4 a 8 c 5 g
ORIGIN

```

```

COMMENT      OS      Homo sapiens (human)
              PN      JP 1999243959-A/8
              PD      14-SEP-1999
              PF      04-MAR-1998 JP 1998052453
              PR      HIDEO YOSHIDA,HIDEKI YANAGI,TAKASHI YURA
              PC      C12N15/09,A61K35/74,A61K35/76,A61K38/00,A61K48/00,C12N15/00,
              PC      A61K37/02
              CC      Strandedness: Double;
              CC      Topology: Linear;
              FH      Key      Location/Qualifiers
              FT      enhancer      1..19.
FEATURES      source
              1..19
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
BASE COUNT    4 a 10 c 4 g 1 t
ORIGIN
Query Match      52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 52.6%; Pred. No. 2.7e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 CCAATNNNNNNNNCCACG 19
      |||||
Db      1 CCAATCGCGCGCACCACG 19

RESULT 9
E28805
LOCUS      E28805      19 bp      DNA      linear      PAT 18-JUN-2001
DEFINITION      Endoplasmic reticulum stress-response regulatory element.
ACCESSION      E28805
VERSION      E28805.1 GI:13020859
KEYWORDS      JP 1999243959-A/9.
SOURCE      Gallus sp.
ORGANISM      Gallus sp.
REFERENCE      Hideo,Y., Hideki,Y. and Takashi,Y.
AUTHORS      Hideo,Y., Hideki,Y. and Takashi,Y.
TITLE      Endoplasmic reticulum stress-response regulatory element
JOURNAL      Patent: JP 1999243959-A 9 14-SEP-1999;
HSP RESEARCH INST INC
COMMENT      OS      Gallus sp. (chicken)
              PN      JP 1999243959-A/9
              PD      14-SEP-1999
              PF      04-MAR-1998 JP 1998052453
              PR      HIDEO YOSHIDA,HIDEKI YANAGI,TAKASHI YURA
              PC      C12N15/09,A61K35/74,A61K35/76,A61K38/00,A61K48/00,C12N15/00,
              PC      A61K37/02
              CC      Strandedness: Double;
              CC      Topology: Linear;
              FH      Key      Location/Qualifiers
              FT      enhancer      1..19.
FEATURES      source
              1..19
              /organism="Gallus sp."
              /mol_type="genomic DNA"
              /db_xref="taxon:9036"
BASE COUNT    5 a 7 c 6 g 1 t
ORIGIN
Query Match      52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 52.6%; Pred. No. 2.7e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 CCAATNNNNNNNNCCACG 19
      |||||
Db      1 CCAATGGGCGCACCACG 19

RESULT 10
E28806
LOCUS      E28806      19 bp      DNA      linear      PAT 18-JUN-2001
DEFINITION      Endoplasmic reticulum stress-response regulatory element.
ACCESSION      E28806
VERSION      E28806.1 GI:13020860
KEYWORDS      JP 1999243959-A/10.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Hideo,Y., Hideki,Y. and Takashi,Y.
AUTHORS      Hideo,Y., Hideki,Y. and Takashi,Y.
TITLE      Endoplasmic reticulum stress-response regulatory element
JOURNAL      Patent: JP 1999243959-A 10 14-SEP-1999;
HSP RESEARCH INST INC
COMMENT      OS      Homo sapiens (human)
              PN      JP 1999243959-A/10
              PD      14-SEP-1999
              PF      04-MAR-1998 JP 1998052453
              PR      HIDEO YOSHIDA,HIDEKI YANAGI,TAKASHI YURA
              PC      C12N15/09,A61K35/74,A61K35/76,A61K38/00,A61K48/00,C12N15/00,
              PC      A61K37/02
              CC      Strandedness: Double;
              CC      Topology: Linear;
              FH      Key      Location/Qualifiers
              FT      enhancer      1..19.
FEATURES      source
              1..19
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
BASE COUNT    6 a 6 c 6 g 1 t
ORIGIN
Query Match      52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 52.6%; Pred. No. 2.7e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 CCAATNNNNNNNNCCACG 19
      |||||
Db      1 CCAATCGGAGCGACCG 19

RESULT 11
E28807
LOCUS      E28807      19 bp      DNA      linear      PAT 18-JUN-2001
DEFINITION      Endoplasmic reticulum stress-response regulatory element.
ACCESSION      E28807
VERSION      E28807.1 GI:13020861
KEYWORDS      JP 1999243959-A/11.
SOURCE      Gallus sp.
ORGANISM      Gallus sp.
REFERENCE      Hideo,Y., Hideki,Y. and Takashi,Y.
AUTHORS      Hideo,Y., Hideki,Y. and Takashi,Y.
TITLE      Endoplasmic reticulum stress-response regulatory element
JOURNAL      Patent: JP 1999243959-A 11 14-SEP-1999;
HSP RESEARCH INST INC
COMMENT      OS      Gallus sp. (chicken)
              PN      JP 1999243959-A/11
              PD      14-SEP-1999
              PF      04-MAR-1998 JP 1998052453
              PR      HIDEO YOSHIDA,HIDEKI YANAGI,TAKASHI YURA
              PC      C12N15/09,A61K35/74,A61K35/76,A61K38/00,A61K48/00,C12N15/00,
              PC      A61K37/02
              CC      Strandedness: Double;

```



Query Match 52.6%; Score 10; DB 6; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.7e+04;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19  
Db 1 CCAATNNNNNNCCACG 19

RESULT 15

E44206 19 bp DNA linear PAT 31-JAN-2002  
LOCUS Endoplasmic reticulum stress transcription factor.  
E44206  
DEFINITION

ACCESSION E44206.1 GI:18633459

VERSION JP 2001054391-A/5.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 19)

Haji,K., Yoshida,H., Mori,K., Yanagi,H. and Yura,T.

Endoplasmic reticulum stress transcription factor

Patent: JP 2001054391-A 5 27-FEB-2001;

HSP RESEARCH INST INC

OS Homo sapiens (human)

PN JP 2001054391-A/5

PD 27-FEB-2001

PF 11-NOV-1999 JP 1999321743

PI KYOSUKE HAJI,HIDEO YOSHIDA,KAZUTOSHI MORI,HIDEKI YANAGI, PI

PC TAKASHI YURA

C12N15/09,C12P21/02//(C12N15/09,C12R1:91),C12N15/00,(C12N15/00, PC

C12R1:91)

CC

FT source

FT source 1. .19 Location/Qualifiers

1. .19 Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 3 a 9 c 5 g 2 t

ORIGIN

Query Match 52.6%; Score 10; DB 6; Length 19;

Best Local Similarity 52.6%; Pred. No. 2.7e+04;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19

Db 1 CCAATCGGCGCTCCACG 19

Search completed: December 4, 2003, 09:21:23

Job time : 1501 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 07:47:59 ; Search time 261 Seconds  
(without alignments)  
196.511 Million cell updates/sec

Title: US-09-606-804-1

Perfect score: 19  
Sequence: 1 ccaatnnnnnnnccacg 19

Scoring table: IDENTITY NUC  
Gapop 10<sup>-7.0</sup>, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq 19Jun03;\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	52.6	19	20	AAZ25631
2	10	52.6	19	20	AAZ25632
3	10	52.6	19	20	AAZ25633
4	10	52.6	19	20	AAZ25634
5	10	52.6	19	20	AAZ25635
6	10	52.6	19	20	AAZ25636
7	10	52.6	19	20	AAZ25637
8	10	52.6	19	20	AAZ25638

9	10	52.6	19	20	AAZ25639	Endoplasmic reticu
10	10	52.6	19	20	AAZ25640	Endoplasmic reticu
11	10	52.6	19	21	AAA28570	GRP78 promoter ERS
12	10	52.6	19	21	AAA28571	GRP78 promoter ERS
13	10	52.6	19	21	AAA28572	GRP78 promoter ERS
14	10	52.6	19	21	AAA28573	GRP94 promoter ERS
15	10	52.6	19	21	AAA28574	GRP94 promoter ERS
16	10	52.6	19	21	AAA28575	GRP94 promoter ERS
17	10	52.6	19	21	AAA28576	GRP94 promoter ERS
18	10	52.6	19	21	AAA28577	GRP94 promoter ERS
19	10	52.6	19	21	AAA28578	Calreticulin promo
20	10	52.6	19	21	AAA28579	Calreticulin promo
21	10	52.6	19	22	AAI70001	ERSE consensus se
22	10	52.6	19	22	AAF28780	ERSE consensus se
23	10	52.6	24	21	AAA28596	Endoplasmic respon
24	10	52.6	24	21	ABQ00200	GRP78 promoter ERS
25	10	52.6	24	24	ABQ04398	Oligonucleotide ad
26	10	52.6	24	24	ABQ04439	Oligonucleotide ad
27	10	52.6	24	24	ABQ10685	Oligonucleotide ad
28	10	52.6	24	24	ABQ10726	Oligonucleotide ad
29	10	52.6	25	19	AAV45529	Oligonucleotide ad
30	10	52.6	25	19	ABQ12221	Helicobacter pylor
31	10	52.6	25	24	ABQ12262	Oligonucleotide ad
32	10	52.6	28	21	AAA28597	Oligonucleotide ad
33	10	52.6	28	21	AAA28599	GRP94 promoter ERS
34	10	52.6	29	21	AAA28602	Calreticulin promo
35	10	52.6	30	21	AAA28598	Tandem repeat comp
36	10	52.6	33	21	AAA28569	GRP94 promoter ERS
37	10	52.6	50	24	ABZ03658	GRP78 promoter tan
38	10	52.6	112	24	ABL41923	Human leukocyte ge
39	10	52.6	122	20	AAZ25660	Nucleotide sequenc
40	10	52.6	122	21	AAZ28566	Human GRP78 promot
41	10	52.6	123	21	ABL80984	5' flanking region
42	10	52.6	126	21	AAA71906	Human ovarian canc
43	10	52.6	127	24	ABL41924	IAK fragment NS fr
44	10	52.6	143	21	AAA71907	Nucleotide sequenc
45	10	52.6	145	22	AAF67488	IAK fragment NS fr
						Novel human polynu

#### ALIGNMENTS

RESULT 1  
AAZ25631  
ID AAZ25631 standard; DNA; 19 BP.  
XX AAZ25631;  
AC AAZ25631;  
XX  
DT 23-DEC-1999 (first entry)  
XX  
DE Endoplasmic reticulum stress competence control element SEQ ID NO:1.  
XX  
KW Endoplasmic reticulum; ER; stress competence; control element;  
KW inhibition; growth; apoptosis; cancer; autoimmune disease;  
KW cystic fibrosis; da.  
XX  
OS Homo sapiens.  
XX  
PN JP1243959-A.  
XX  
PD 14-SEP-1999.  
XX  
PF 04-MAR-1998; 98JP-0052453.  
XX  
PR 04-MAR-1998; 98JP-0052453.  
XX  
PA (HSPK-) HSP KENKYUSHO KK.  
XX  
DR WPI; 1999-603708/52.  
XX  
PT New control element for stress competence of endoplasmic reticulum -  
PT useful for inhibition of growth and induction of apoptosis in cancer  
PT cells

```
XX PS Claim 1; Page 10; 25pp; Japanese.
XX CC The present invention specifically claims an element shown by: (A) a
CC 19 bp base sequence, CCAATNNNN NNNCCACG (ERSE); or (B) a modified base
CC sequence having replaced 1-3 bases with the other base(s), which induces
CC transcription with stress on endoplasmic reticulum used for stress
CC competence of endoplasmic reticulum. Also described are: (1) a DNA having
CC transcription inducing activity with stress on endoplasmic reticulum
CC containing the above mentioned element, optionally further containing a
CC promoter DNA; and (2) a vector containing the element optionally with the
CC DNA. The element can be used for the inhibition of growth and induction
CC of apoptosis of cancer cells, and improvement of symptoms of autoimmune
CC diseases and cystic fibrosis by inhibition of autoantibody formation.
XX
SQ Sequence 19 BP; 3 A; 5 C; 1 G; 1 T; 9 other;

Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 1 CCAATNNNNNNNNCCACG 19

RESULT 2
AAZ25632
ID .AAZ25632 standard; DNA; 19 BP.
XX
AC AAZ25632;
XX
DT 23-DEC-1999 (first entry)
XX
DE Endoplasmic reticulum stress competence control element SEQ ID NO:5.
XX
KW Endoplasmic reticulum; ER; stress competence; control element;
KW inhibition; growth; apoptosis; cancer; autoimmune disease;
KW cystic fibrosis; ds.
XX
OS Homo sapiens.
XX
PN JP11243959-A.
XX
PD 14-SEP-1999.
XX
PF 04-MAR-1998; 98JP-0052453.
XX
PR 04-MAR-1998; 98JP-0052453.
XX
PA (HSPK-) HSP KENKYUSHO KK.
XX
DR WPI; 1999-603708/52.
XX
PT New control element for stress competence of endoplasmic reticulum -
PT useful for inhibition of growth and induction of apoptosis in cancer
PT cells
XX
PS Example 1; Fig 3; 25pp; Japanese.
XX
CC The present invention specifically claims an element shown by: (A) a
CC 19 bp base sequence, CCAATNNNN NNNCCACG (ERSE); or (B) a modified base
CC sequence having replaced 1-3 bases with the other base(s), which induces
CC transcription with stress on endoplasmic reticulum used for stress
CC competence of endoplasmic reticulum. Also described are: (1) a DNA
CC having transcription inducing activity with stress on endoplasmic
CC reticulum containing the above mentioned element, optionally further
CC containing a promoter DNA; and (2) a vector containing the element
CC optionally with the DNA. The element can be used for the inhibition of
CC growth and induction of apoptosis of cancer cells, and improvement of
CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
CC autoantibody formation. AAZ25632 to AAZ25657 represent elements used in
CC an example from the present invention.
XX
SQ Sequence 19 BP; 3 A; 5 C; 1 G; 1 T; 9 other;

Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 1 CCAATNNNNNNNNCCACG 19

RESULT 3
AAZ25633
ID AAZ25633 standard; DNA; 19 BP.
XX
AC AAZ25633;
XX
DT 23-DEC-1999 (first entry)
XX
DE Endoplasmic reticulum stress competence control element SEQ ID NO:6.
XX
KW Endoplasmic reticulum; ER; stress competence; control element;
KW inhibition; growth; apoptosis; cancer; autoimmune disease;
KW cystic fibrosis; ds.
XX
OS Mus sp.
XX
PN JP11243959-A.
XX
PD 14-SEP-1999.
XX
PF 04-MAR-1998; 98JP-0052453.
XX
PR 04-MAR-1998; 98JP-0052453.
XX
PA (HSPK-) HSP KENKYUSHO KK.
XX
DR WPI; 1999-603708/52.
XX
PT New control element for stress competence of endoplasmic reticulum -
PT useful for inhibition of growth and induction of apoptosis in cancer
PT cells
XX
PS Example 1; Fig 3; 25pp; Japanese.
XX
CC The present invention specifically claims an element shown by: (A) a
CC 19 bp base sequence, CCAATNNNN NNNCCACG (ERSE); or (B) a modified base
CC sequence having replaced 1-3 bases with the other base(s), which induces
CC transcription with stress on endoplasmic reticulum used for stress
CC competence of endoplasmic reticulum. Also described are: (1) a DNA
CC having transcription inducing activity with stress on endoplasmic
CC reticulum containing the above mentioned element, optionally further
CC containing a promoter DNA; and (2) a vector containing the element
CC optionally with the DNA. The element can be used for the inhibition of
CC growth and induction of apoptosis of cancer cells, and improvement of
CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
CC autoantibody formation. AAZ25632 to AAZ25657 represent elements used in
CC an example from the present invention.
XX
SQ Sequence 19 BP; 4 A; 8 C; 5 G; 2 T; 0 other;

Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 1 CCAATCGGAGGCGCTCCACG 19

RESULT 4
AAZ25634
```



```

ID AA225634 standard; DNA; 19 BP.
XX AC AA225634;
XX DT 23-DEC-1999 (first entry)
XX DE Endoplasmic reticulum stress competence control element SEQ ID NO:7.
XX KW Endoplasmic reticulum; ER; stress competence; control element;
XX KW inhibition; growth; apoptosis; cancer; autoimmune disease;
XX KW cystic fibrosis; ds.
XX OS Rattus sp.
XX PN JP11243959-A.
XX PD 14-SEP-1999.
XX PF 04-MAR-1998; 98JP-0052453.
XX PR 04-MAR-1998; 98JP-0052453.
XX PA (HSPK-) HSP KENKYUSHO KK.
XX DR WPI; 1999-603708/52.
XX PT New control element for stress competence of endoplasmic reticulum -
XX PT useful for inhibition of growth and induction of apoptosis in cancer
XX PS cells
XX PS Example 1; Fig 3; 25pp; Japanese.
XX CC The present invention specifically claims an element shown by: (A) a
XX CC 19 bp base sequence, CCAATNNNNN NNNCCACG (ERSE); or (B) a modified base
XX CC sequence having replaced 1-3 bases with the other base(s), which induces
XX CC transcription with stress on endoplasmic reticulum used for stress
XX CC competence of endoplasmic reticulum. Also described are: (1) a DNA
XX CC having transcription inducing activity with stress on endoplasmic
XX CC reticulum containing the above mentioned element, optionally further
XX CC containing a promoter DNA; and (2) a vector containing the element
XX CC optionally with the DNA. The element can be used for the inhibition of
XX CC growth and induction of apoptosis of cancer cells, and improvement of
XX CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
XX CC autoantibody formation. AA225632 to AA225657 represent elements used in
XX CC an example from the present invention.
XX SQ Sequence 19 BP; 4 A; 8 C; 5 G; 2 T; 0 other;

Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNN NNNCCACG 19
Db 1 CCAATCGGAGCTCCACG 19

RESULT 5
AA225635
ID AA225635 standard; DNA; 19 BP.
XX AC AA225635;
XX DT 23-DEC-1999 (first entry)
XX DE Endoplasmic reticulum stress competence control element SEQ ID NO:8.
XX KW Endoplasmic reticulum; ER; stress competence; control element;
XX KW inhibition; growth; apoptosis; cancer; autoimmune disease;
XX KW cystic fibrosis; ds.
XX OS Homo sapiens.
XX PT New control element for stress competence of endoplasmic reticulum -
XX PT useful for inhibition of growth and induction of apoptosis in cancer

```

---

```

PN JP11243959-A.
XX PD 14-SEP-1999.
XX PF 04-MAR-1998; 98JP-0052453.
XX PR 04-MAR-1998; 98JP-0052453.
XX PA (HSPK-) HSP KENKYUSHO KK.
XX DR WPI; 1999-603708/52.
XX PT New control element for stress competence of endoplasmic reticulum -
XX PT useful for inhibition of growth and induction of apoptosis in cancer
XX PS cells
XX PS Example 1; Fig 3; 25pp; Japanese.
XX CC The present invention specifically claims an element shown by: (A) a
XX CC 19 bp base sequence, CCAATNNNNN NNNCCACG (ERSE); or (B) a modified base
XX CC sequence having replaced 1-3 bases with the other base(s), which induces
XX CC transcription with stress on endoplasmic reticulum used for stress
XX CC competence of endoplasmic reticulum. Also described are: (1) a DNA
XX CC having transcription inducing activity with stress on endoplasmic
XX CC reticulum containing the above mentioned element, optionally further
XX CC containing a promoter DNA; and (2) a vector containing the element
XX CC optionally with the DNA. The element can be used for the inhibition of
XX CC growth and induction of apoptosis of cancer cells, and improvement of
XX CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
XX CC autoantibody formation. AA225632 to AA225657 represent elements used in
XX CC an example from the present invention.
XX SQ Sequence 19 BP; 4 A; 10 C; 4 G; 1 T; 0 other;

Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNN NNNCCACG 19
Db 1 CCAATCGGCGGCACCACG 19

RESULT 6
AA225636
ID AA225636 standard; DNA; 19 BP.
XX AC AA225636;
XX DT 23-DEC-1999 (first entry)
XX DE Endoplasmic reticulum stress competence control element SEQ ID NO:9.
XX KW Endoplasmic reticulum; ER; stress competence; control element;
XX KW inhibition; growth; apoptosis; cancer; autoimmune disease;
XX KW cystic fibrosis; ds.
XX OS Gallus sp.
XX PN JP11243959-A.
XX PD 14-SEP-1999.
XX PF 04-MAR-1998; 98JP-0052453.
XX PR 04-MAR-1998; 98JP-0052453.
XX PA (HSPK-) HSP KENKYUSHO KK.
XX DR WPI; 1999-603708/52.
XX PT New control element for stress competence of endoplasmic reticulum -
XX PT useful for inhibition of growth and induction of apoptosis in cancer

```

```

PT cells
XX Example 1; Fig 3; 25pp; Japanese.
PS
XX The present invention specifically claims an element shown by: (A) a
CC 19 bp base sequence, CCAATNNNN NNNCCACG (ERSE); or (B) a modified base
CC sequence having replaced 1-3 bases with the other base(s), which induces
CC transcriptions with stress on endoplasmic reticulum used for stress
CC competence of endoplasmic reticulum. Also described are: (1) a DNA
CC having transcription inducing activity with stress on endoplasmic
CC reticulum containing the above mentioned element, optionally further
CC containing a promoter DNA; and (2) a vector containing the element
CC optionally with the DNA. The element can be used for the inhibition of
CC growth and induction of apoptosis of cancer cells, and improvement of
CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
CC autoantibody formation. AA225632 to AA225657 represent elements used in
CC an example from the present invention.
XX
SQ Sequence 19 BP; 5 A; 7 C; 6 G; 1 T; 0 other;
Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 CCAATNNNNNNNNCCACG 19
Db 1 CCAATGGGAGCGCACCG 19
RESULT 7
AA225637
ID AA225637 standard; DNA; 19 BP.
XX
AC AA225637;
XX
DT 23-DEC-1999 (first entry)
XX
DE Endoplasmic reticulum stress competence control element SEQ ID NO:10.
XX
KW Endoplasmic reticulum; ER; stress competence; control element;
KW inhibition; growth; apoptosis; cancer; autoimmune disease;
KW cystic fibrosis; ds.
XX
OS Homo sapiens.
XX
PN JP11243959-A.
XX
PD 14-SEP-1999.
XX
PF 04-MAR-1998; 98JP-0052453.
XX
PR 04-MAR-1998; 98JP-0052453.
XX
PA (HSPK-) HSP KENKYUSHO KK.
XX
DR WPI; 1999-603708/52.
XX
PT New control element for stress competence of endoplasmic reticulum -
PT useful for inhibition of growth and induction of apoptosis in cancer
PT cells
PS Example 1; Fig 3; 25pp; Japanese.
XX
CC The present invention specifically claims an element shown by: (A) a
CC 19 bp base sequence, CCAATNNNN NNNCCACG (ERSE); or (B) a modified base
CC sequence having replaced 1-3 bases with the other base(s), which induces
CC transcriptions with stress on endoplasmic reticulum used for stress
CC competence of endoplasmic reticulum. Also described are: (1) a DNA
CC having transcription inducing activity with stress on endoplasmic
CC reticulum containing the above mentioned element, optionally further
CC containing a promoter DNA; and (2) a vector containing the element
CC optionally with the DNA. The element can be used for the inhibition of
CC growth and induction of apoptosis of cancer cells, and improvement of
CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
CC autoantibody formation. AA225632 to AA225657 represent elements used in
CC an example from the present invention.
XX
SQ Sequence 19 BP; 5 A; 7 C; 6 G; 1 T; 0 other;
Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 CCAATNNNNNNNNCCACG 19
Db 1 CCAATGGGAGCGCACCG 19
RESULT 8
AA225638
ID AA225638 standard; DNA; 19 BP.
XX
AC AA225638;
XX
DT 23-DEC-1999 (first entry)
XX
DE Endoplasmic reticulum stress competence control element SEQ ID NO:11.
XX
KW Endoplasmic reticulum; ER; stress competence; control element;
KW inhibition; growth; apoptosis; cancer; autoimmune disease;
KW cystic fibrosis; ds.
XX
OS Gallus sp.
XX
PN JP11243959-A.
XX
PD 14-SEP-1999.
XX
PF 04-MAR-1998; 98JP-0052453.
XX
PR 04-MAR-1998; 98JP-0052453.
XX
PA (HSPK-) HSP KENKYUSHO KK.
XX
DR WPI; 1999-603708/52.
XX
PT New control element for stress competence of endoplasmic reticulum -
PT useful for inhibition of growth and induction of apoptosis in cancer
PT cells
PS Example 1; Fig 3; 25pp; Japanese.
XX
CC The present invention specifically claims an element shown by: (A) a
CC 19 bp base sequence, CCAATNNNN NNNCCACG (ERSE); or (B) a modified base
CC sequence having replaced 1-3 bases with the other base(s), which induces
CC transcriptions with stress on endoplasmic reticulum used for stress
CC competence of endoplasmic reticulum. Also described are: (1) a DNA
CC having transcription inducing activity with stress on endoplasmic
CC reticulum containing the above mentioned element, optionally further
CC containing a promoter DNA; and (2) a vector containing the element
CC optionally with the DNA. The element can be used for the inhibition of
CC growth and induction of apoptosis of cancer cells, and improvement of
CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
CC autoantibody formation. AA225632 to AA225657 represent elements used in
CC an example from the present invention.
XX
SQ Sequence 19 BP; 4 A; 9 C; 5 G; 1 T; 0 other;
Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 CCAATNNNNNNNNCCACG 19
Db 1 CCAATCGACGCGCGCACG 19

```

```

RESULT 9
AAZ25639
ID AAZ25639 standard; DNA; 19 BP.
XX
XX AC AAZ25639;
XX
XX DT 23-DEC-1999 (first entry)
XX
XX DE Endoplasmic reticulum stress competence control element SEQ ID NO:12.
XX
XX KW Endoplasmic reticulum; ER; stress competence; control element;
XX KW inhibition; growth; apoptosis; cancer; autoimmune disease;
XX KW cystic fibrosis; ds.
XX
XX OS Homo sapiens.
XX
XX PN JP11243959-A.
XX
XX PD 14-SEP-1999.
XX
XX PF 04-MAR-1998; 98JJP-0052453.
XX
XX PR 04-MAR-1998; 98JJP-0052453.
XX
XX PA (HSPK-) HSP KENKYUSHO KK.
XX
XX DR WPI; 1999-603708/52.
XX
XX PT New control element for stress competence of endoplasmic reticulum -
XX PT useful for inhibition of growth and induction of apoptosis in cancer
XX PT cells
XX
XX PS Example 1; Fig 3; 25pp; Japanese.
XX
XX CC The present invention specifically claims an element shown by: (A) a
XX CC 19 bp base sequence, CCAATNNNNN NNNCCACG (ERSE); or (B) a modified base
XX CC sequence having replaced 1-3 bases with the other base(s), which induces
XX CC transcription with stress on endoplasmic reticulum used for stress
XX CC competence of endoplasmic reticulum. Also described are: (1) a DNA
XX CC having transcription inducing activity with stress on endoplasmic
XX CC reticulum containing the above mentioned element, optionally further
XX CC containing a promoter DNA; and (2) a vector containing the element
XX CC optionally with the DNA. The element can be used for the inhibition of
XX CC growth and induction of apoptosis of cancer cells, and improvement of
XX CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
XX CC autoantibody formation. AAZ25632 to AAZ25657 represent elements used in
XX CC an example from the present invention.
XX
XX SQ Sequence 19 BP; 5 A; 6 C; 5 G; 3 T; 0 other;

Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNN NNNCCACG 19
    |||||
Db 1 CCAATGATGTCGACACG 19
    |||||

RESULT 10
AAZ25640
ID AAZ25640 standard; DNA; 19 BP.
XX
XX AC AAZ25640;
XX
XX DT 23-DEC-1999 (first entry)
XX
XX DE Endoplasmic reticulum stress competence control element SEQ ID NO:13.
XX
XX KW Endoplasmic reticulum; ER; stress competence; control element;
XX KW inhibition; growth; apoptosis; cancer; autoimmune disease;
XX KW cystic fibrosis; ds.

```

```

XX Mus sp.
XX
XX PN JP11243959-A.
XX
XX PD 14-SEP-1999.
XX
XX PF 04-MAR-1998; 98JJP-0052453.
XX
XX PR 04-MAR-1998; 98JJP-0052453.
XX
XX PA (HSPK-) HSP KENKYUSHO KK.
XX
XX DR WPI; 1999-603708/52.
XX
XX PT New control element for stress competence of endoplasmic reticulum -
XX PT useful for inhibition of growth and induction of apoptosis in cancer
XX PT cells
XX
XX PS Example 1; Fig 3; 25pp; Japanese.
XX
XX CC The present invention specifically claims an element shown by: (A) a
XX CC 19 bp base sequence, CCAATNNNNN NNNCCACG (ERSE); or (B) a modified base
XX CC sequence having replaced 1-3 bases with the other base(s), which induces
XX CC transcription with stress on endoplasmic reticulum used for stress
XX CC competence of endoplasmic reticulum. Also described are: (1) a DNA
XX CC having transcription inducing activity with stress on endoplasmic
XX CC reticulum containing the above mentioned element, optionally further
XX CC containing a promoter DNA; and (2) a vector containing the element
XX CC optionally with the DNA. The element can be used for the inhibition of
XX CC growth and induction of apoptosis of cancer cells, and improvement of
XX CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
XX CC autoantibody formation. AAZ25632 to AAZ25657 represent elements used in
XX CC an example from the present invention.
XX
XX SQ Sequence 19 BP; 5 A; 6 C; 6 G; 2 T; 0 other;

Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNN NNNCCACG 19
    |||||
Db 1 CCAATGATGTCGACACG 19
    |||||

RESULT 11
AAA28570
ID AAA28570 standard; DNA; 19 BP.
XX
XX AC AAA28570;
XX
XX DT 29-AUG-2000 (first entry)
XX
XX DE GRP78 promoter ERSE1-like sequence.
XX
XX KW Endoplasmic reticulum; stress; ER; transcription factor;
XX KW transcription; regulatory element; ERSE; bZIP; chaperone; treatment;
XX KW prophylaxis; cancer; arteriosclerosis; ischaemia; wound healing;
XX KW cystic fibrosis; ulcer; gene therapy; recombinant gene; human;
XX KW gene expression; GRP; glucose regulated protein; promoter; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200029429-A2.
XX
XX PD 25-MAY-2000.
XX
XX PF 12-NOV-1999; 99WO-JP06305.
XX
XX KW 13-NOV-1998; 98JP-0324227.
XX PR 09-JUN-1999; 99JP-0163112.
XX

```

PA (HSPR-) HSP RES INST INC.  
 PI Haze K, Yoshida H, Mori K, Yanagi H, Yura T;  
 XX WPI; 2000-387736/33.  
 XX  
 PT New endoplasmic reticulum stress transcription factor (known as bZIP)  
 PT for controlling expression of endoplasmic reticulum chaperone, useful  
 PT for treating cancers, arteriosclerosis, cystic fibrosis, ischemic  
 PT diseases, wounds and ulcers  
 PS  
 PS Example 1; Fig 3; 157pp; English.  
 XX  
 CC An endoplasmic reticulum stress transcription factor (bZIP)  
 CC capable of regulating transcription inducing activity exhibited by an  
 CC element (ERSE) can be used in a method for controlling expression of  
 CC an endoplasmic reticulum chaperone. The method comprises expressing  
 CC bZIP. The method can be used for expression of a foreign protein by  
 CC positively regulating expression of an endoplasmic reticulum  
 CC chaperone gene. bZIP is useful for controlling the expression of  
 CC endoplasmic reticulum chaperone either positively or negatively in  
 CC cells and therefore is useful for treatment or prophylaxis of  
 CC cancers, arteriosclerosis, cystic fibrosis, ischaemic diseases,  
 CC wounds and ulcers. bZIP also maintains the correct conformation of  
 CC the endoplasmic reticulum chaperone and thereby increases the  
 CC expression of a foreign protein. This sequence taken from the  
 CC glucose regulating protein (GRP) promoter GRP78 contains an ERSE like  
 CC sequence.  
 XX  
 SQ Sequence 19 BP; 3 A; 9 C; 5 G; 2 T; 0 other;  
 Query Match 52.6%; Score 10; DB 21; Length 19;  
 Best Local Similarity 52.6%; Pred. No. 1.8e+03;  
 Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 OY 1 CCAATNNNNNNNNCCACG 19  
 Db ||||| |||||  
 1 CCAATCGGCGCTCCACG 19  
 RESULT 12  
 AAA28571  
 ID AAA28571 standard; DNA; 19 BP.  
 XX  
 AC AAA28571;  
 XX  
 DT 29-AUG-2000 (first entry)  
 XX  
 DE GRP78 promoter ERSE1-like sequence.  
 XX  
 KW Endoplasmic reticulum; stress; ER; transcription factor;  
 KW transcription; regulatory element; ERSE; bZIP; chaperone; treatment;  
 KW prophylaxis; cancer; arteriosclerosis; ischaemia; wound healing;  
 KW cystic fibrosis; ulcer; gene therapy; recombinant gene; mouse;  
 KW gene expression; GRP; glucose regulated protein; promoter; ss.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200029429-A2.  
 XX  
 XX 25-MAY-2000.  
 XX  
 PF 12-NOV-1999; 99WO-JP06305.  
 XX  
 PR 13-NOV-1998; 98JP-0324227.  
 PR 09-JUN-1999; 99JP-0163112.  
 XX  
 XX (HSPR-) HSP RES INST INC.  
 XX  
 PI Haze K, Yoshida H, Mori K, Yanagi H, Yura T;  
 XX WPI; 2000-387736/33.  
 XX

PT New endoplasmic reticulum stress transcription factor (known as bZIP)  
 PT for controlling expression of endoplasmic reticulum chaperone, useful  
 PT for treating cancers, arteriosclerosis, cystic fibrosis, ischemic  
 PT diseases, wounds and ulcers  
 PS  
 PS Example 1; Fig 3; 157pp; English.  
 XX  
 CC An endoplasmic reticulum stress transcription factor (bZIP)  
 CC capable of regulating transcription inducing activity exhibited by an  
 CC element (ERSE) can be used in a method for controlling expression of  
 CC an endoplasmic reticulum chaperone. The method comprises expressing  
 CC bZIP. The method can be used for expression of a foreign protein by  
 CC positively regulating expression of an endoplasmic reticulum  
 CC chaperone gene. bZIP is useful for controlling the expression of  
 CC endoplasmic reticulum chaperone either positively or negatively in  
 CC cells and therefore is useful for treatment or prophylaxis of  
 CC cancers, arteriosclerosis, cystic fibrosis, ischaemic diseases,  
 CC wounds and ulcers. bZIP also maintains the correct conformation of  
 CC the endoplasmic reticulum chaperone and thereby increases the  
 CC expression of a foreign protein. This sequence taken from the  
 CC glucose regulating protein (GRP) promoter GRP78 contains an ERSE like  
 CC sequence.  
 XX  
 SQ Sequence 19 BP; 4 A; 8 C; 5 G; 2 T; 0 other;  
 Query Match 52.6%; Score 10; DB 21; Length 19;  
 Best Local Similarity 52.6%; Pred. No. 1.8e+03;  
 Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 OY 1 CCAATNNNNNNNNCCACG 19  
 Db ||||| |||||  
 1 CCAATCGGAGGCTCCACG 19  
 RESULT 13  
 AAA28572  
 ID AAA28572 standard; DNA; 19 BP.  
 XX  
 AC AAA28572;  
 XX  
 DT 29-AUG-2000 (first entry)  
 XX  
 DE GRP78 promoter ERSE1-like sequence.  
 XX  
 KW Endoplasmic reticulum; stress; ER; transcription factor;  
 KW transcription; regulatory element; ERSE; bZIP; chaperone; treatment;  
 KW prophylaxis; cancer; arteriosclerosis; ischaemia; wound healing;  
 KW cystic fibrosis; ulcer; gene therapy; recombinant gene; rat;  
 KW gene expression; GRP; glucose regulated protein; promoter; ss.  
 XX  
 OS Rattus rattus.  
 XX  
 PN WO200029429-A2.  
 XX  
 XX 25-MAY-2000.  
 XX  
 PF 12-NOV-1999; 99WO-JP06305.  
 XX  
 PR 13-NOV-1998; 98JP-0324227.  
 PR 09-JUN-1999; 99JP-0163112.  
 XX  
 XX (HSPR-) HSP RES INST INC.  
 XX  
 PI Haze K, Yoshida H, Mori K, Yanagi H, Yura T;  
 XX WPI; 2000-387736/33.  
 XX  
 PT New endoplasmic reticulum stress transcription factor (known as bZIP)  
 PT for controlling expression of endoplasmic reticulum chaperone, useful  
 PT for treating cancers, arteriosclerosis, cystic fibrosis, ischemic  
 PT diseases, wounds and ulcers  
 PS  
 PS Example 1; Fig 3; 157pp; English.

```

XX CC An endoplasmic reticulum stress transcription factor (bZIP)
CC capable of regulating transcription inducing activity exhibited by an
CC element (ERSE) can be used in a method for controlling expression of
CC an endoplasmic reticulum chaperone. The method comprises expressing
CC bZIP. The method can be used for expression of a foreign protein by
CC positively regulating expression of an endoplasmic reticulum
CC chaperone gene. bZIP is useful for controlling the expression of
CC endoplasmic reticulum chaperone either positively or negatively in
CC cells and therefore is useful for treatment or prophylaxis of
CC cancers, arteriosclerosis, cystic fibrosis, ischaemic diseases,
CC wounds and ulcers. bZIP also maintains the correct conformation of
CC the endoplasmic reticulum chaperone and thereby increases the
CC expression of a foreign protein. This sequence taken from the
CC glucose regulating protein (GRP) promoter GRP78 contains an ERSE like
CC sequence.
XX SQ Sequence 19 BP; 4 A; 8 C; 5 G; 2 T; 0 other;
      Query Match      52.6%; Score 10; DB 21; Length 19;
      Best Local Similarity 52.6%; Pred. No. 1.8e+03;
      Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
   |||||
Db 1 CCAATCGAGCGCTCCACG 19
   |||||

RESULT 14
AAA28573
ID AAA28573 standard; DNA; 19 BP.
AC AAA28573;
XX 29-AUG-2000 (first entry)
DE GRP94 promoter ERSE1-like sequence.
XX Endoplasmic reticulum; stress; ER; transcription factor;
KW transcription; regulatory element; ERSE; bZIP; chaperone; treatment;
KW prophylaxis; cancer; arteriosclerosis; ischaemia; wound healing;
KW cystic fibrosis; ulcer; gene therapy; recombinant gene; human;
KW gene expression; GRP; glucose regulated protein; promoter; ss.
XX Homo sapiens.
OS
XX WO200029429-A2.
PN
XX 25-MAY-2000.
PD
XX 12-NOV-1999; 99WO-JP06305.
PF
XX 13-NOV-1998; 98JP-0324227.
PR
XX 09-JUN-1999; 99JP-0163112.
PX
XX (HSPR-) HSP RES INST INC.
PA
XX Haze K, Yoshida H, Mori K, Yanagi H, Yura T;
PI WPI; 2000-387736/33.
PT
XX New endoplasmic reticulum stress transcription factor (known as bZIP)
PT for controlling expression of endoplasmic reticulum chaperone, useful
PT for treating cancers, arteriosclerosis, cystic fibrosis, ischemic
PT diseases, wounds and ulcers
XX Example 1; Fig 3; 157pp; English.
PS
XX An endoplasmic reticulum stress transcription factor (bZIP)
CC capable of regulating transcription inducing activity exhibited by an
CC element (ERSE) can be used in a method for controlling expression of
CC an endoplasmic reticulum chaperone. The method comprises expressing
CC bZIP. The method can be used for expression of a foreign protein by
XX positively regulating expression of an endoplasmic reticulum
XX chaperone gene. bZIP is useful for controlling the expression of
XX cells and therefore is useful for treatment or prophylaxis of
XX cancers, arteriosclerosis, cystic fibrosis, ischaemic diseases,
XX wounds and ulcers. bZIP also maintains the correct conformation of

```

```

CC positively regulating expression of an endoplasmic reticulum
CC chaperone gene. bZIP is useful for controlling the expression of
CC endoplasmic reticulum chaperone either positively or negatively in
CC cells and therefore is useful for treatment or prophylaxis of
CC cancers, arteriosclerosis, cystic fibrosis, ischaemic diseases,
CC wounds and ulcers. bZIP also maintains the correct conformation of
CC the endoplasmic reticulum chaperone and thereby increases the
CC expression of a foreign protein. This sequence taken from the
CC glucose regulating protein (GRP) promoter GRP94 contains an ERSE like
CC sequence.
XX SQ Sequence 19 BP; 4 A; 10 C; 4 G; 1 T; 0 other;
      Query Match      52.6%; Score 10; DB 21; Length 19;
      Best Local Similarity 52.6%; Pred. No. 1.8e+03;
      Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
   |||||
Db 1 CCAATCGCGCGCACACG 19
   |||||

RESULT 15
AAA28574
ID AAA28574 standard; DNA; 19 BP.
AC AAA28574;
XX 29-AUG-2000 (first entry)
DE GRP94 promoter ERSE1-like sequence.
XX Endoplasmic reticulum; stress; ER; transcription factor;
KW transcription; regulatory element; ERSE; bZIP; chaperone; treatment;
KW prophylaxis; cancer; arteriosclerosis; ischaemia; wound healing;
KW cystic fibrosis; ulcer; gene therapy; recombinant gene; chicken;
KW gene expression; GRP; glucose regulated protein; promoter; ss.
XX Gallus domesticus.
OS
XX WO200029429-A2.
PN
XX 25-MAY-2000.
PD
XX 12-NOV-1999; 99WO-JP06305.
PF
XX 13-NOV-1998; 98JP-0324227.
PR
XX 09-JUN-1999; 99JP-0163112.
PX
XX (HSPR-) HSP RES INST INC.
PA
XX Haze K, Yoshida H, Mori K, Yanagi H, Yura T;
PI WPI; 2000-387736/33.
PT
XX New endoplasmic reticulum stress transcription factor (known as bZIP)
PT for controlling expression of endoplasmic reticulum chaperone, useful
PT for treating cancers, arteriosclerosis, cystic fibrosis, ischemic
PT diseases, wounds and ulcers
XX Example 1; Fig 3; 157pp; English.
PS
XX An endoplasmic reticulum stress transcription factor (bZIP)
CC capable of regulating transcription inducing activity exhibited by an
CC element (ERSE) can be used in a method for controlling expression of
CC an endoplasmic reticulum chaperone. The method comprises expressing
CC bZIP. The method can be used for expression of a foreign protein by
XX positively regulating expression of an endoplasmic reticulum
XX chaperone gene. bZIP is useful for controlling the expression of
XX cells and therefore is useful for treatment or prophylaxis of
XX cancers, arteriosclerosis, cystic fibrosis, ischaemic diseases,
XX wounds and ulcers. bZIP also maintains the correct conformation of

```

CC the endoplasmic reticulum chaperone and thereby increases the  
CC expression of a foreign protein. This sequence taken from the  
CC glucose regulating protein (GRP) promoter GRP94 contains an ERSE like  
CC sequence.  
XX  
SQ Sequence 19 BP; 5 A; 7 C; 6 G; 1 T; 0 other;  
Query Match 52.6%; Score 10; DB 21; Length 19;  
Best Local Similarity 52.6%; Pred. No. 1.8e+03;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
|||  
Db 1 CCAATGGGAGCGCACG 19  
|||

Search completed: December 4, 2003, 08:56:09  
Job time : 262 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 08:47:17 ; Search time 75 Seconds  
(without alignments)  
111.817 Million cell updates/sec

Title: US-09-606-804-1

Perfect score: 19

Sequence: 1 ccaatnnnnnnnnccacg 19

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	52.6	19	4	US-09-813-937-1
2	10	52.6	203	4	US-09-313-294A-1696
3	10	52.6	261	4	US-09-389-681-206
4	10	52.6	261	4	US-09-620-405B-206
5	10	52.6	261	4	US-09-339-338-206
6	10	52.6	261	4	US-09-433-826B-206
7	10	52.6	261	4	US-09-604-287A-206
8	10	52.6	270	4	US-09-313-294A-890
9	10	52.6	272	1	US-08-248-474-36
10	10	52.6	272	3	US-08-756-849-36
11	10	52.6	282	4	US-09-252-991A-987
12	10	52.6	285	4	US-09-252-991A-60
13	10	52.6	296	2	US-08-716-942-20
14	10	52.6	296	4	US-09-130-337A-20
15	10	52.6	300	4	US-09-313-294A-7348
16	10	52.6	360	6	5196523-1
17	10	52.6	387	3	US-08-804-372A-23
18	10	52.6	392	4	US-09-404-879A-372
19	10	52.6	399	6	5196523-23
20	10	52.6	421	6	5196523-22
21	10	52.6	424	1	US-08-609-657-15
22	10	52.6	432	4	US-09-252-991A-9959
23	10	52.6	457	4	US-09-702-705-861
24	10	52.6	457	4	US-09-736-457-861
25	10	52.6	467	3	US-09-257-584-8
26	10	52.6	511	4	US-09-702-705-1138
27	10	52.6	511	4	US-09-736-457-1138

C	28	10	52.6	522	4	US-09-663-600A-149	Sequence 149, App
	29	10	52.6	531	4	US-09-222-575-123	Sequence 123, App
	30	10	52.6	531	4	US-09-389-681-123	Sequence 123, App
	31	10	52.6	531	4	US-09-620-405B-123	Sequence 123, App
	32	10	52.6	531	4	US-09-339-338-123	Sequence 123, App
	33	10	52.6	531	4	US-09-433-826B-123	Sequence 123, App
	34	10	52.6	531	4	US-09-604-287A-123	Sequence 123, App
	35	10	52.6	584	4	US-09-663-600A-55	Sequence 55, Appl
	36	10	52.6	641	4	US-09-573-906-7	Sequence 7, Appl
	37	10	52.6	708	4	US-09-252-991A-3364	Sequence 3364, Ap
	38	10	52.6	754	3	US-09-020-956-20	Sequence 20, Appl
	39	10	52.6	754	3	US-09-030-607-20	Sequence 20, Appl
	40	10	52.6	754	4	US-09-439-313-20	Sequence 20, Appl
	41	10	52.6	754	4	US-09-352-616A-20	Sequence 20, Appl
	42	10	52.6	754	4	US-09-232-149A-20	Sequence 20, Appl
	43	10	52.6	770	4	US-09-573-906-6	Sequence 6, Appl
	44	10	52.6	843	4	US-09-171-209-38	Sequence 38, Appl
	45	10	52.6	849	3	US-08-998-416-552	Sequence 552, App

## ALIGNMENTS

## RESULT 1

US-09-813-937-1  
; Sequence 1, Application US/09813937  
; Patent No. 6506600  
; GENERAL INFORMATION:  
; APPLICANT: HERMONAT, Paul L.  
; APPLICANT: MAINE, Michael  
; APPLICANT: LIU, Yong  
; TITLE OF INVENTION: SECRETING PRODUCTS FROM SKIN BY ADENO-ASSOCIATED VIRUS (AAV) GEN.  
; TITLE OF INVENTION: TRANSFER  
; FILE REFERENCE: 023533/0116  
; CURRENT APPLICATION NUMBER: US/09/813.937  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: US 60/191092  
; FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 19  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: glucose response element  
; NAME/KEY: misc feature  
; LOCATION: (6)-(14)  
; OTHER INFORMATION: "N" at positions 6 - 14 can be A, C, G or T  
US-09-813-937-1

Query Match 52.6%, Score 10; DB 4; Length 19;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
| | | | | | | | | | | | | | | | | |  
Db 1 CCAATNNNNNNNNCCACG 19

## RESULT 2

US-09-313-294A-1696/c  
; Sequence 1696, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313.294A  
; CURRENT FILING DATE: 1999-05-14

```
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 1696
; LENGTH: 205
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700551354H1
US-09-313-294A-1696

Query Match      52.6%; Score 10; DB 4; Length 261;
Best Local Similarity 52.6%; Pred. No. 5.2e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 122 CCAATCAGACAGACCCACG 104

RESULT 3
US-09-389-681-206
; Sequence 206, Application US/09389681A
; Patent No. 6518237
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C3
; CURRENT APPLICATION NUMBER: US/09/389,681A
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-389-681-206

Query Match      52.6%; Score 10; DB 4; Length 261;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 140 CCAATCTTCATCTCCACG 158

RESULT 4
US-09-620-405B-206
; Sequence 206, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-620-405B-206

Query Match      52.6%; Score 10; DB 4; Length 261;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 140 CCAATCTTCATCTCCACG 158

RESULT 5
US-09-339-338-206
; Sequence 206, Application US/09339338A
; Patent No. 6573368
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C2
; CURRENT APPLICATION NUMBER: US/09/339,338A
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-339-338-206

Query Match      52.6%; Score 10; DB 4; Length 261;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 140 CCAATCTTCATCTCCACG 158

RESULT 6
US-09-433-826B-206
; Sequence 206, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-433-826B-206

Query Match      52.6%; Score 10; DB 4; Length 261;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 140 CCAATCTTCATCTCCACG 158
```



```
RESULT 7
US-09-604-287A-206
; Sequence 206, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-604-287A-206

Query Match          52.6%; Score 10; DB 4; Length 261;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 140 CCAATTCCTCATCTCCACG 158

RESULT 8
US-09-313-294A-890/c
; Sequence 890, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 890
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700550018H1
; NAME/KEY: unsure
; LOCATION: 213-215, 218-221, 232-233, 236-237, 239-240, 242, 249, 251, 253, 261,
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-890

Query Match          52.6%; Score 10; DB 4; Length 270;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 121 CCAATCAGACAGACCCACG 103

RESULT 9
US-08-248-474-36
; Sequence 36, Application US/08248474
; Patent No. 5612471
```

```
; GENERAL INFORMATION:
; APPLICANT: McK. Bird, David
; APPLICANT: WILSON, Mark A.
; TITLE OF INVENTION: NEMATODE-INDUCE GENES IN TOMATO
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,474
; FILING DATE: 25-MAY-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 2307E-535
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Lycopersicon esculentum cv 'Rutgers Large
; ORGANISM: Red'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..272
; OTHER INFORMATION: /standard_name= "DB# 155"
US-08-248-474-36

Query Match          52.6%; Score 10; DB 1; Length 272;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 97 CCAATATTATTACCACG 115

RESULT 10
US-08-756-849-36
; Sequence 36, Application US/08756849
; Patent No. 6093810
; GENERAL INFORMATION:
; APPLICANT: Bird, David McK.
; APPLICANT: Wilson, Mark A.
; TITLE OF INVENTION: Nematode-Induced Genes in Tomato
; NUMBER OF SEQUENCES: 129
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/756,849  
; FILING DATE: 26-NOV-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/248,474  
; FILING DATE: 25-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 023070-053510US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 272 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: Lycopersicon esculentum cv 'Rutgers Large Red'  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1..272  
; OTHER INFORMATION: /standard\_name= "DB# 155"  
; US-08-756-849-36

Query Match 52.6%; Score 10; DB 3; Length 272;  
Best Local Similarity 52.6%; Pred. No. 5.4e+02;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
Db 97 CCAATATTATTTTACCACG 115

RESULT 11  
US-09-252-991A-987  
; Sequence 987, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Terragen Diversity Inc.  
; APPLICANT: Radomski, Christopher C. A.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 987  
; LENGTH: 282  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-987

Query Match 52.6%; Score 10; DB 4; Length 282;  
Best Local Similarity 52.6%; Pred. No. 5.4e+02;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
Db 49 CCAATGCTGATCCCCACG 67

RESULT 12  
US-09-252-991A-60  
; Sequence 60, Application US/09252991A

; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 60  
; LENGTH: 285  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-60

Query Match 52.6%; Score 10; DB 4; Length 285;  
Best Local Similarity 52.6%; Pred. No. 5.5e+02;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
Db 224 CCAATGACCGATACCACG 242

## RESULT 13

US-08-716-942-20/c  
; Sequence 20, Application US/08716942  
; Patent No. 5849491  
; GENERAL INFORMATION:  
; APPLICANT: Terragen Diversity Inc.  
; APPLICANT: Radomski, Christopher C. A.  
; APPLICANT: Seow, Kah Tong  
; APPLICANT: Warren, R. Antony J.  
; APPLICANT: Yap, Wai Ho  
; TITLE OF INVENTION: METHOD FOR ISOLATING XYLANASE GENE  
; TITLE OF INVENTION: SEQUENCES FROM SOIL DNA, COMPOSITIONS USEFUL IN SUCH METHOD AN  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Oppedahl & Larson  
; STREET: 1992 Commerce Street, Suite 309  
; CITY: Yorktown Heights  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10598-4412  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS 5.0  
; SOFTWARE: Word Perfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/716,942  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/004,157  
; FILING DATE: 20-Sep-95  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Marina T. Larson  
; REGISTRATION NUMBER: 32,038  
; REFERENCE/DOCKET NUMBER: TERR.P-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (914) 245-3252  
; TELEFAX: (914) 962-4330  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 296  
; TYPE: nucleic acid

```
; STRANDEDNESS: DOUBLE
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM:
; FEATURE:
; NAME/KEY: fragment of xylanase gene from degenerate primer
; NAME/KEY: amplification of soil DNA
US-08-716-942-20

Query Match      52.6%; Score 10; DB 2; Length 296;
Best Local Similarity 52.6%; Pred. No. 5.5e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
Db 217 CCAATGGTTGTGCCACG 199

RESULT 14
US-09-130-337A-20/c
; Sequence 20, Application US/09130337A
; Patent No. 6441148
; GENERAL INFORMATION:
; APPLICANT: Radomski, CCA
; APPLICANT: Seow, KT
; APPLICANT: Warren, RAJ
; APPLICANT: Yap, WH
; TITLE OF INVENTION: METHOD FOR AMPLIFICATION OF XYLANASE GENE FRAGMENTS
; TITLE OF INVENTION: FROM SOIL DNA, COMPOSITIONS USEFUL IN SUCH METHOD AND COMPOSITIO
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 9993-004
; CURRENT APPLICATION NUMBER: US/09/130,337A
; CURRENT FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 08/716,942
; PRIOR FILING DATE: 1996-09-20
; PRIOR APPLICATION NUMBER: 60/004,157
; PRIOR FILING DATE: 1995-09-22
; NUMBER OF SEQ ID NOS: 25
; SEQ ID NO 20
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Unknown organism
; FEATURE:
; OTHER INFORMATION: Description of unknown organism: soil microbe
US-09-130-337A-20

Query Match      52.6%; Score 10; DB 4; Length 296;
Best Local Similarity 52.6%; Pred. No. 5.5e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
Db 217 CCAATGGTTGTGCCACG 199

RESULT 15
US-09-313-294A-7348/c
; Sequence 7348, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
```

```
; SEQ ID NO 7348
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700381587H1
; NAME/KEY: unsure
; LOCATION: 18, 228, 295
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-7348

Query Match      52.6%; Score 10; DB 4; Length 300;
Best Local Similarity 52.6%; Pred. No. 5.5e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
Db 292 CCAATCGTACCTCCACG 274

Search completed: December 4, 2003, 09:58:51
Job time : 79 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 09:21:30 ; Search time 254 Seconds  
(without alignment)  
248.616 Million cell updates/sec

Title: US-09-606-804-1

Perfect score: 19

Sequence: 1 ccaatnnnnnnnccacg 19

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	52.6	19	14	US-10-340-759-1
C 2	10	52.6	24	11	Sequence 1, Appli
C 3	10	52.6	25	9	Sequence 207, App
C 4	10	52.6	25	9	Sequence 35, Appl
C 5	10	52.6	25	11	US-09-940-185-207
C 6	10	52.6	100	10	US-09-940-185-4188
C 7	10	52.6	123	10	US-09-969-373-1152
C 8	10	52.6	136	12	US-09-867-701-3962
C 9	10	52.6	136	12	US-09-922-449B-18
C 10	10	52.6	136	12	US-09-922-449B-19
C 11	10	52.6	143	12	US-09-922-449B-8
C 12	10	52.6	149	12	US-09-922-449B-24
C 13	10	52.6	150	12	US-09-922-449B-25
C 14	10	52.6	167	14	US-09-922-449B-10
C 15	10	52.6	210	10	US-10-001-883-5
C 16	10	52.6	214	9	US-09-974-300-8096
					Sequence 8096, Ap
					Sequence 59, Appl
					Sequence 59, Appl

C 17	10	52.6	214	13	US-10-010-742-59	Sequence 59, Appl
C 18	10	52.6	219	10	US-09-783-590-10136	Sequence 10136, A
C 19	10	52.6	225	14	US-10-060-036-3284	Sequence 3284, Ap
C 20	10	52.6	237	9	US-09-777-564-1013	Sequence 1013, Ap
C 21	10	52.6	237	14	US-10-015-219-1013	Sequence 1013, Ap
C 22	10	52.6	238	10	US-09-960-352-13155	Sequence 13155, A
C 23	10	52.6	247	10	US-09-878-574-6181	Sequence 6181, Ap
C 24	10	52.6	261	9	US-09-604-387A-206	Sequence 206, App
C 25	10	52.6	261	10	US-09-339-338-206	Sequence 206, App
C 26	10	52.6	261	11	US-09-551-621-206	Sequence 206, App
C 27	10	52.6	261	12	US-10-124-805-206	Sequence 206, App
C 28	10	52.6	261	13	US-10-007-805-206	Sequence 206, App
C 29	10	52.6	261	14	US-10-076-622-206	Sequence 206, App
C 30	10	52.6	262	12	US-10-099-926-1836	Sequence 1836, Ap
C 31	10	52.6	262	13	US-10-033-528-1836	Sequence 1836, Ap
C 32	10	52.6	267	9	US-09-294-093B-1896	Sequence 1896, Ap
C 33	10	52.6	268	9	US-09-923-876-5857	Sequence 5857, Ap
C 34	10	52.6	271	10	US-09-783-590-11139	Sequence 11139, A
C 35	10	52.6	272	11	US-09-835-976B-57	Sequence 57, Appl
C 36	10	52.6	274	10	US-09-867-701-5278	Sequence 5278, Ap
C 37	10	52.6	280	10	US-09-878-574-246	Sequence 246, App
C 38	10	52.6	283	9	US-09-294-093B-3259	Sequence 3259, Ap
C 39	10	52.6	293	9	US-09-294-093B-6179	Sequence 6179, Ap
C 40	10	52.6	301	11	US-09-918-995-24371	Sequence 24371, A
C 41	10	52.6	306	10	US-09-974-300-8100	Sequence 8100, Ap
C 42	10	52.6	339	10	US-09-878-574-1186	Sequence 1186, Ap
C 43	10	52.6	348	11	US-09-899-495-18	Sequence 18, Appl
C 44	10	52.6	351	10	US-09-938-842A-1451	Sequence 1451, Ap
C 45	10	52.6	351	10	US-09-796-692-6226	Sequence 6226, Ap

ALIGNMENTS

RESULT 1  
US-10-340-759-1

; Sequence 1, Application US/10340759  
; Publication No. US20030104575A1

; GENERAL INFORMATION:

; APPLICANT: HERMONAT, Paul L.

; APPLICANT: MANE, Michael

; APPLICANT: LIU, Yong

; TITLE OF INVENTION: SECRETING PRODUCTS FROM SKIN BY ADENO-ASSOCIATED VIRUS (AAV) GENI

; TITLE OF INVENTION: TRANSFER

; FILE REFERENCE: 023533/0116

; CURRENT APPLICATION NUMBER: US/10/340,759

; PRIOR FILING DATE: 2003-01-13

; PRIOR APPLICATION NUMBER: US/09/813,937A

; PRIOR FILING DATE: 2001-03-22

; PRIOR APPLICATION NUMBER: US 60/191092

; PRIOR FILING DATE: 2000-03-22

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 19

; TYPE: DNA

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: glucose response element

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (6)..(14)

; OTHER INFORMATION: "N" at positions 6 - 14 can be A, C, G or T

US-10-340-759-1

Query Match 52.6%; Score 10; DB 14; Length 19;

Best Local Similarity 100.0%; Pred No. 4.1e+03;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNCCACG 19

Db 1 CCAATNNNNNNNCCACG 19

```
RESULT 2
US-09-940-185-207/c
; Sequence 207, Application US/09940185
; Publication No. US20030096239A1
; GENERAL INFORMATION:
; APPLICANT: Gunderson, Kevin
; APPLICANT: Chee, Mark
; TITLE OF INVENTION: Probes and Decoder Oligonucleotides
; FILE REFERENCE: A-69605-1
; CURRENT APPLICATION NUMBER: US 60/227,948
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/227,948
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/228,854
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 4768
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 207
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Computer Generated Probe Sequence.
US-09-940-185-207

Query Match          52.6%; Score 10; DB 11; Length 24;
Best Local Similarity 52.6%; Pred. No. 4.2e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 23 CCAATATTACGTGACCACG 5

RESULT 3
US-09-402-100-35/c
; Sequence 35, Application US/09402100
; Patent No. US20010019834A1
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co, LTD
; APPLICANT: Kim, Byung-O
; APPLICANT: Shin, Sung-Seup
; APPLICANT: Yu, Young-Hyo
; APPLICANT: Park, Myung-Hwan
; APPLICANT: Choi, Deok-Joon
; APPLICANT: Jung, Hyung-Jin
; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Helicobacter
; FILE REFERENCE: 0136/OG140
; CURRENT APPLICATION NUMBER: US/09/402,100
; PRIOR FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: KR 97-11950
; EARLIER FILING DATE: 1997-03-31
; EARLIER APPLICATION NUMBER: KR 97-11951
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1..7)
; OTHER INFORMATION: Oligonucleotide
US-09-402-100-35

Query Match          52.6%; Score 10; DB 9; Length 25;
Best Local Similarity 52.6%; Pred. No. 4.2e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 23 CCAATATTACGTGACCACG 5

RESULT 4
US-09-940-185-4188/c
; Sequence 4188, Application US/09940185
; Publication No. US20030096239A1
; GENERAL INFORMATION:
; APPLICANT: Gunderson, Kevin
; APPLICANT: Chee, Mark
; TITLE OF INVENTION: Probes and Decoder Oligonucleotides
; FILE REFERENCE: A-69605-1
; CURRENT APPLICATION NUMBER: US/09/940,185
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/227,948
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/228,854
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 4768
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4188
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Computer Generated Probe Sequence.
US-09-940-185-4188

Query Match          52.6%; Score 10; DB 11; Length 25;
Best Local Similarity 52.6%; Pred. No. 4.2e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 24 CCAATATTACGTGACCACG 6

RESULT 5
US-09-969-373-1152/c
; Sequence 1152, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
; APPLICANT: Effertz, Roger J.
; APPLICANT: Hauge, Brian M.
; TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
; FILE REFERENCE: 38-10(52679)A
; CURRENT APPLICATION NUMBER: US/09/969,373
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 09/754,853
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 09/760,427
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 09/855,768
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 1152
; LENGTH: 100
; TYPE: DNA
; ORGANISM: Glycine max
US-09-969-373-1152

Query Match          52.6%; Score 10; DB 10; Length 100;
Best Local Similarity 52.6%; Pred. No. 4.9e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 97 CCAATCATTTCATCCACG 79

RESULT 6
US-09-867-701-3962
; Sequence 3962, Application US/09867701
```

Db 20 CCAATTTTACCATCCACG 2

; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3962  
; LENGTH: 123  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-867-701-3962

Query Match 52.6%; Score 10; DB 10; Length 123;  
Best Local Similarity 52.6%; Pred. No. 5e+03;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 4 CCAATCTGTCGGACCACG 22

## RESULT 7

US-09-922-449B-18  
; Sequence 18, Application US/09922449B  
; Publication No. US20030148278A1  
; GENERAL INFORMATION:  
; APPLICANT: BioInside Gesellschaft fur Biodiagnostik, Auftragsforschung und Consulting  
; APPLICANT: mbH  
; TITLE OF INVENTION: Test kit and method for quantitatively detecting genetically modified  
; FILE REFERENCE: 101215-68  
; CURRENT APPLICATION NUMBER: US/09/922,449B  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: PCT/EP00/009835  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: DE 199 06 169.6  
; PRIOR FILING DATE: 1999-02-08  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 18  
; LENGTH: 136  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: target IAC DNA for the Bt-176  
US-09-922-449B-18

Query Match 52.6%; Score 10; DB 12; Length 136;  
Best Local Similarity 52.6%; Pred. No. 5.1e+03;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 27 CCAATTTTCGCTCCACG 45

## RESULT 8

US-09-922-449B-19  
; Sequence 19, Application US/09922449B  
; Publication No. US20030148278A1  
; GENERAL INFORMATION:  
; APPLICANT: BioInside Gesellschaft fur Biodiagnostik, Auftragsforschung und Consulting  
; APPLICANT: mbH  
; TITLE OF INVENTION: Test kit and method for quantitatively detecting genetically modified  
; FILE REFERENCE: 101215-68  
; CURRENT APPLICATION NUMBER: US/09/922,449B  
; CURRENT FILING DATE: 2001-08-03

; CURRENT APPLICATION NUMBER: US/09/922,449B  
; CURRENT FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: PCT/EP00/009835  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: DE 199 06 169.6  
; PRIOR FILING DATE: 1999-02-08  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 19  
; LENGTH: 136  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: reference IAC DNA for the Bt-176  
US-09-922-449B-19

Query Match 52.6%; Score 10; DB 12; Length 136;  
Best Local Similarity 52.6%; Pred. No. 5.1e+03;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 27 CCAATTTTCGCTCCACG 45

## RESULT 9

US-09-922-449B-8  
; Sequence 8, Application US/09922449B  
; Publication No. US20030148278A1  
; GENERAL INFORMATION:  
; APPLICANT: BioInside Gesellschaft fur Biodiagnostik, Auftragsforschung und Consulting  
; APPLICANT: mbH  
; TITLE OF INVENTION: Test kit and method for quantitatively detecting genetically modified  
; FILE REFERENCE: 101215-68  
; CURRENT APPLICATION NUMBER: US/09/922,449B  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: PCT/EP00/009835  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: DE 199 06 169.6  
; PRIOR FILING DATE: 1999-02-08  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 142  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: target IAC DNA for the RRS 9e  
US-09-922-449B-8

Query Match 52.6%; Score 10; DB 12; Length 142;  
Best Local Similarity 52.6%; Pred. No. 5.1e+03;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 31 CCAATTTTCGCTCCACG 49

## RESULT 10

US-09-922-449B-24  
; Sequence 24, Application US/09922449B  
; Publication No. US20030148278A1  
; GENERAL INFORMATION:  
; APPLICANT: BioInside Gesellschaft fur Biodiagnostik, Auftragsforschung und Consulting  
; APPLICANT: mbH  
; TITLE OF INVENTION: Test kit and method for quantitatively detecting genetically modified  
; FILE REFERENCE: 101215-68  
; CURRENT APPLICATION NUMBER: US/09/922,449B  
; CURRENT FILING DATE: 2001-08-03

; PRIOR APPLICATION NUMBER: PCT/EP00/009835  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: DE 199 06 169.6  
; PRIOR FILING DATE: 1999-02-08  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 24  
; LENGTH: 143  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: target IAC DNA for the RRS ge  
US-09-922-449B-24

Query Match 52.6%; Score 10; DB 12; Length 143;

Best Local Similarity 52.6%; Pred. No. 5.1e+03;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 31 CCAATTTTCGCTCCACG 49

## RESULT 11

US-09-922-449B-25  
; Sequence 25, Application US/09922449B  
; Publication No. US20030148278A1  
; GENERAL INFORMATION:  
; APPLICANT: BioInside Gesellschaft fur Bodiagnostik, Auftragsforschung und Consulting  
; APPLICANT: mbH  
; TITLE OF INVENTION: Test kit and method for quantitatively detecting genetically modified  
; FILE OF INVENTION: in foodstuff by means of fluorescent-coupled PCR  
; FILE REFERENCE: 101215-68  
; CURRENT APPLICATION NUMBER: US/09/922,449B  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: PCT/EP00/009835  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: DE 199 06 169.6  
; PRIOR FILING DATE: 1999-02-08  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 25  
; LENGTH: 149  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: reference IAC DNA for the RRS  
US-09-922-449B-25

Query Match 52.6%; Score 10; DB 12; Length 149;

Best Local Similarity 52.6%; Pred. No. 5.1e+03;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 31 CCAATTTTCGCTCCACG 49

## RESULT 12

US-09-922-449B-10  
; Sequence 10, Application US/09922449B  
; Publication No. US20030148278A1  
; GENERAL INFORMATION:  
; APPLICANT: BioInside Gesellschaft fur Bodiagnostik, Auftragsforschung und Consulting  
; APPLICANT: mbH  
; TITLE OF INVENTION: Test kit and method for quantitatively detecting genetically modified  
; FILE OF INVENTION: in foodstuff by means of fluorescent-coupled PCR  
; FILE REFERENCE: 101215-68  
; CURRENT APPLICATION NUMBER: US/09/922,449B  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: PCT/EP00/009835  
; PRIOR FILING DATE: 2000-02-07

; PRIOR APPLICATION NUMBER: DE 199 06 169.6  
; PRIOR FILING DATE: 1999-02-08  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 150  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: reference IAC DNA for the RRS  
US-09-922-449B-10

Query Match 52.6%; Score 10; DB 12; Length 150;

Best Local Similarity 52.6%; Pred. No. 5.1e+03;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 32 CCAATTTTCGCTCCACG 50

## RESULT 13

US-10-001-883-5/c  
; Sequence 5, Application US/10001883  
; Publication No. US20030022188A1  
; GENERAL INFORMATION:  
; APPLICANT: Macina, Roberto  
; APPLICANT: Recipon, Herve  
; APPLICANT: Pluta, Jason  
; APPLICANT: Ghosh, Malavika  
; APPLICANT: Sun, Yongming  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Pro  
; FILE REFERENCE: DEX-0271  
; CURRENT APPLICATION NUMBER: US/10/001,883  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 60/252,059  
; PRIOR FILING DATE: 2000-11-20  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 167  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-001-883-5

Query Match 52.6%; Score 10; DB 14; Length 167;

Best Local Similarity 52.6%; Pred. No. 5.2e+03;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 160 CCAATTTTGGAAATCACCACG 142

## RESULT 14

US-09-974-300-8096/c  
; Sequence 8096, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; APPLICANT: Clausen, Ib Groth  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; FILE OF INVENTION: Expression  
; FILE REFERENCE: 10085.500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; CURRENT FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/279,526  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 8481

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8096
; LENGTH: 210
; TYPE: DNA
; ORGANISM: Bacillus clausii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(210)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-8096

Query Match          52.6%; Score 10; DB 10; Length 210;
Best Local Similarity 52.6%; Pred. No. 5.3e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy      1 CCAATNNNNNNCCACG 19
        |||||
Db      152 CCAATTCGCTGTCACG 134

RESULT 15
US-09-778-320-59/c
; Sequence 59, Application US/09778320
; Patent No. US20010034052A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Day, Craig H.
; APPLICANT: Jiang, Yugu
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Wang, Tongrong
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.491CS
; CURRENT APPLICATION NUMBER: US/09/778,320
; CURRENT FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 214
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(214)
; OTHER INFORMATION: n = A,T,C or G
US-09-778-320-59

Query Match          52.6%; Score 10; DB 9; Length 214;
Best Local Similarity 52.6%; Pred. No. 5.3e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy      1 CCAATNNNNNNCCACG 19
        |||||
Db      88 CCAATTCGCTGTCACG 70

Search completed: December 4, 2003, 10:57:37
Job time : 257 secs
```



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 05:58:56 ; Search time 2068 Seconds  
(without alignments)  
223.300 Million cell updates/sec

Title: US-09-606-804-1

Perfect score: 19  
Sequence: 1 ccaatnnnnnnnnccacg 19

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	52.6	43	28	BH857113
C 2	10	52.6	47	29	BX292617
C 3	10	52.6	58	29	AL949322
4	10	52.6	76	9	AA475936

5	10	52.6	79	13	B0756881
C 6	10	52.6	80	9	AL897168
C 7	10	52.6	80	28	BH251489
8	10	52.6	82	13	BQ099298
C 9	10	52.6	86	28	BH232183
C 10	10	52.6	97	28	BH222106
C 11	10	52.6	99	14	CB338675
C 12	10	52.6	99	28	AZ921866
C 13	10	52.6	100	10	BF086094
14	10	52.6	101	28	BH583741
C 15	10	52.6	103	9	AA927925
C 16	10	52.6	109	28	BH864991
17	10	52.6	110	10	BF094257
18	10	52.6	110	28	AZ720239
19	10	52.6	112	29	BZ662888
C 20	10	52.6	113	28	BH222899
21	10	52.6	114	29	CC458196
22	10	52.6	115	9	A1938545
23	10	52.6	115	10	AW945325
C 24	10	52.6	115	12	BG950375
25	10	52.6	115	13	BW091136
26	10	52.6	115	29	CC029017
27	10	52.6	115	29	AG024826
28	10	52.6	116	29	CC037007
C 29	10	52.6	119	9	A1549504
C 30	10	52.6	120	28	BH230343
31	10	52.6	120	28	BH905244
C 32	10	52.6	121	14	CB097190
C 33	10	52.6	121	28	BH222096
34	10	52.6	123	9	AA405167
35	10	52.6	124	28	AQ073902
C 36	10	52.6	124	28	BH812753
C 37	10	52.6	124	28	BH847929
C 38	10	52.6	129	28	AZ919927
39	10	52.6	130	28	BH895687
40	10	52.6	130	29	BZ663067
41	10	52.6	131	29	BZ386507
42	10	52.6	132	10	BF361951
43	10	52.6	132	10	BE936064
44	10	52.6	132	28	BH753521
45	10	52.6	132	28	BH854781

## ALIGNMENTS

RESULT 1

BH857113

LOCUS

DEFINITION

BH857113 43 bp DNA linear GSS 08-JUL-2002  
SALK\_076821.44.95.x Arabidopsis thaliana TDNA insertion lines  
Arabidopsis thaliana genomic clone SALK\_076821.44.95.x, genomic  
survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

i eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu  
 This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated exon of At4g29100.  
 Class: TDNA tagged.

#### FEATURES

##### Location/Qualifiers

```
1. .43
  /organism="Arabidopsis thaliana"
  /mol_type="genomic DNA"
  /strain="Columbia 0"
  /db_xref="taxon:3702"
  /clone="SALK_076821.44.95.x"
  /clone_lib="Arabidopsis thaliana TDNA insertion lines"
  /note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"
```

BASE COUNT 11 a 11 c 10 g 11 t  
 ORIGIN

Query Match 52.6%; Score 10; DB 28; Length 43;  
 Best Local Similarity 52.6%; Pred. No. 2e+04;  
 Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNCCACG 19  
 ||||| |||||  
 Db 18 CCAATGAATTCGCCACG 36

#### RESULT 2

BX292617/c 47 bp DNA linear GSS 07-MAR-2003  
 LOCUS  
 DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-455E08-018778,  
 genomic survey sequence.

ACCESSION BX292617  
 VERSION BX292617.1 GI:28891613  
 KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 Strizhov, N., Li, Y., Rosso, M., Viehoever, P., Dekker, K., Saedler, H.  
 and Weisshaar, B.  
 A pipeline for automated high-throughput generation of FSTs  
 (flanking sequence tags) from Arabidopsis thaliana T-DNA  
 transformed lines  
 Unpublished

REFERENCE 2 Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weisshaar, B.  
 A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)  
 for flanking sequence tag based reverse genetics  
 Unpublished

REFERENCE 3 (bases 1 to 47)

Li, Y., Strizhov, N., Rosso, M. and Weisshaar, B.  
 Direct Submission  
 Submitted (07-MAR-2003) Weisshaar B., Max-Planck-Institut fuer  
 Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany  
 This sequence is recovered from the left border of the T-DNA. It  
 indicates an insertion within the locus defined by clone f1913. The  
 sequences are generated at the MPI for Plant Breeding Research in  
 the context of the GABI-Kat project. GABI-Kat is part of the German  
 Plant Genomics program designated 'GABI'. Information on line  
 availability can be found at:  
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

Location/Qualifiers

```
1. .47
  /organism="Arabidopsis thaliana"
  /mol_type="genomic DNA"
  /strain="Columbia 0"
  /db_xref="taxon:3702"
```

#### FEATURES

##### source

```
/clone="CK-455E08-018778"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana
plants (T1) which were transformed with the T-DNA from
vector pAC161. The lines contain one or more T-DNA
insertions. The DNA fragment(s) resulting from the PCR
were directly sequenced to determine the genomic sequence
flanking the insertion. Sequences displaying significant
similarity to the A. thaliana nuclear genome sequence were
processed for submission. T-DNA derived sequences were
removed"
```

BASE COUNT 17 a 4 c 10 g 16 t  
 ORIGIN

Query Match 52.6%; Score 10; DB 29; Length 47;  
 Best Local Similarity 52.6%; Pred. No. 2.1e+04;  
 Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNCCACG 19  
 ||||| |||||

Db 22 CCAATAAAACTCTCCACG 4

#### RESULT 3

AL949322/c

LOCUS

DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-319H03-015860,  
 genomic survey sequence.

ACCESSION AL949322

VERSION AL949322.1 GI:24405944

KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 Strizhov, N., Li, Y., Rosso, M., Viehoever, P., Dekker, K., Saedler, H.  
 and Weisshaar, B.

A pipeline for automated high-throughput generation of FSTs  
 (flanking sequence tags) from Arabidopsis thaliana T-DNA  
 transformed lines  
 Unpublished

REFERENCE 2 Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weisshaar, B.  
 A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)  
 for flanking sequence tag based reverse genetics  
 Unpublished

REFERENCE 3 (bases 1 to 58)

Rosso, M., Strizhov, N., Li, Y. and Weisshaar, B.  
 Direct Submission

Submitted (21-OCT-2002) Weisshaar B., Max-Planck-Institut fuer  
 Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany  
 This sequence is recovered from the left border of the T-DNA. It  
 indicates an insertion close to or within gene Atlg50730. The  
 sequences are generated at the MPI for Plant Breeding Research in  
 the context of the GABI-Kat project. GABI-Kat is part of the German  
 Plant Genomics program designated 'GABI'. Information on line  
 availability can be found at:  
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

Location/Qualifiers

```
1. .58
  /organism="Arabidopsis thaliana"
  /mol_type="genomic DNA"
  /strain="Columbia 0"
  /db_xref="taxon:3702"
  /clone="GK-319H03-015860"
  /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
  /note="PCR was performed on DNA from Arabidopsis thaliana
plants (T1) which were transformed with the T-DNA from
vector pAC161. The lines contain one or more T-DNA
insertions. The DNA fragment(s) resulting from the PCR
were directly sequenced to determine the genomic sequence
```

#### FEATURES

##### source

flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

```

BASE COUNT      21 a  11 c  10 g  16 t
ORIGIN
Query Match      52.6%; Score 10; DB 29; Length 58;
Best Local Similarity 52.6%; Pred. No. 2.3e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNCCACG 19
    |||||
Db 34 CCAATCTGTGTAGCCACG 16
    |||||

```

```

RESULT 4
AA475936
LOCUS
DEFINITION
VH25B10.t1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone
IMAGE:876475 5' similar to TR:E196749 E196749 MRNA; EXPRESSED
SEQUENCE TAG ;, mRNA sequence.

```

```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
EST.

```

```

REFERENCE
AUTHORS
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

```

```

TITLE
JOURNAL
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:515955

```

Trace considered overall poor quality  
Possible reversed clone: similarity on wrong strand  
Seq primer: -28m13 rev2 ET from Amersham  
High quality sequence stop: 1.

#### FEATURES

```

Location/Qualifiers
1..76
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:876475"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares_mammary_gland NbMMG"

```

```

/notes="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
) with a modified polylinker; Site 1: Not I; Site 2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5',
TGTACCAATCTGAAGTGGAGCGGCCGCAATGGTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima

```

```

BASE COUNT      19 a  28 c  16 g  13 t
ORIGIN
Query Match      52.6%; Score 10; DB 9; Length 76;
Best Local Similarity 52.6%; Pred. No. 2.5e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy 1 CCAATNNNNNNNCCACG 19
    |||||
Db 40 CCAATGTTTGTAGCCACG 58
    |||||

```

```

RESULT 5
BQ756881
LOCUS
DEFINITION
EBem09_SQ005_E15_R embryo, 1 Day germination, no treatment, cv
Optic, EBem09 Hordeum vulgare subsp. vulgare cDNA clone
EBem09_SQ005_E15 5', mRNA sequence.

```

```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare subsp. vulgare

```

```

REFERENCE
AUTHORS
Hadley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardie,L.,
Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
Development of Barley Transcriptome Resources
Unpublished
Contact: Waugh R, Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: est@scri.sari.ac.uk.

```

```

FEATURES
Location/Qualifiers
1..79
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Optic"
/db_xref="taxon:112509"
/clone="EBem09_SQ005_E15"
/tissue_type="embryo"
/dev_stage="1 Day Germination"
/lab_host="DH10B"
/clone_lib="embryo, 1 Day germination, no treatment, cv
Optic, EBem09"
/notes="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from embryos dissected from germinating grains (1
day) in glasshouse grown barley plants. Developed as part
of the barley transcriptome resources of BBSRC/SEERAD
funded cereal 1G (Investigating Gene Function) project."

```

```

BASE COUNT      19 a  18 c  17 g  25 t
ORIGIN
Query Match      52.6%; Score 10; DB 13; Length 79;
Best Local Similarity 52.6%; Pred. No. 2.6e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy 1 CCAATNNNNNNNCCACG 19
    |||||
Db 42 CCAATTTTGTCAACACG 60
    |||||

```

```

RESULT 6
AL897168/c
LOCUS
DEFINITION
AL897168 XGC-egg Silurana tropicalis cDNA clone TEGg01ld06 3', mRNA
80 bp mRNA linear EST 16-SEP-2002

```



revealed the presence of an additional Trichostrongyloidea cattle nematode, *Cooperia oncophora*. Sequences in this library may derive from either *Ostertagia* or *Cooperia*. The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. The cDNA was made by using Dynabead oligo-dT priming (Dyna). PCR based library using a modified protocol from the SMART PCR cDNA Synthesis Kit from Clontech. Directionally cloned into the UDG sites of pAMPI. Nematodes were provided by Dr. Louis Gasbarre of the USDA, Beltsville, MD (lgasbarr@nri.barc.usda.gov). Third stage exsheathed larvae were collected from 14 day fecal-sphagnum moss cultures of *Ostertagia* eggs. The larvae were recovered by overnight passage on a Baermann apparatus, and then cleaned by passage through a 20 micron nylon mesh. The larvae were then subjected to a treatment with 1.25% chlorox to induce excystation. The larvae were washed with 5 changes of PBS and then pelleted and snap frozen in liquid nitrogen."

BASE COUNT 24 a 24 c 12 g 22 t  
ORIGIN

Query Match 52.6%; Score 10; DB 13; Length 82;  
Best Local Similarity 52.6%; Pred. No. 2.6e+04;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 12 CCAATGGAGAGTGTCACG 30

RESULT 9  
BH232183/c  
LOCUS  
DEFINITION 86 bp DNA linear GSS 08-NOV-2001  
1006166D12.V1 1006 - RescueMu Grid G Zea mays genomic, genomic survey sequence.

ACCESSION  
VERSION BH232183.1 GI:16837119  
KEYWORDS  
SOURCE GSS.  
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 86)  
Walbot, V.  
Maize genomic sequences found using engineered RescueMu transposon  
Unpublished  
Contact: Walbot V

Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Possible ligation site so sequence was trimmed. Post-ligation sequence submitted separately.  
Plate: 1006166 row: 23  
Class: transposon-tagged.  
Location/Qualifiers

FEATURES  
source  
1..86  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="mixed background W23/A188/B73"  
/db\_xref="taxon:4577"  
/tissue\_type="leaf"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="1006 - RescueMu Grid G"  
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA."

Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu'. Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT 13 a 24 c 37 g 12 t  
ORIGIN

Query Match 52.6%; Score 10; DB 28; Length 86;  
Best Local Similarity 52.6%; Pred. No. 2.6e+04;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 61 CCAATCATTTGACGCCACG 43

RESULT 10  
BH222106/c  
LOCUS  
DEFINITION 97 bp DNA linear GSS 08-NOV-2001  
1006105C10.X1 1006 - RescueMu Grid G Zea mays genomic, genomic survey sequence.

ACCESSION  
VERSION BH222106.1 GI:16816606  
KEYWORDS  
SOURCE GSS.  
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 97)  
Walbot, V.  
Maize genomic sequences found using engineered RescueMu transposon  
Unpublished  
Contact: Walbot V

Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 1006105 row: 13  
Class: transposon-tagged.  
Location/Qualifiers

FEATURES  
source  
1..97

/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="mixed background W23/A188/B73"  
/db\_xref="taxon:4577"  
/tissue\_type="leaf"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="1006 - RescueMu Grid G"  
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu'. Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT 11 a 27 c 44 g 15 t  
ORIGIN

Query Match 52.6%; Score 10; DB 28; Length 97;  
Best Local Similarity 52.6%; Pred. No. 2.8e+04;

```

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19
    |||||
Db 92 CCAATTGAGCACCACG 74

RESULT 11
CB338675/c
LOCUS
DEFINITION
  CB338675 99 bp mRNA linear EST 13-MAR-2003
  kb59804.y1 Brugia malayi L3 pAMP1 v2 Brugia malayi cDNA 5' similar
  to TR:O16159 O16159 CYSTATIN-TYPE CYSTEINE PROTEINASE INHIBITOR. ;,
  mRNA sequence.
ACCESSION
  CB338675
VERSION
  CB338675.1 GI:28941376
KEYWORDS
  EST.
SOURCE
  Brugia malayi
  Brugia malayi
  Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
  Onchocercidae; Brugia.
REFERENCE
  1 (bases 1 to 99)
  McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.,
  Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,
  Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Teagareishvili,R.,
  Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
  M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
  Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
  Wilson,R.
  The Washington Univ. Nematode EST Project, 1999
  Unpublished
  Contact: McCarter JP
  The Washington Univ. Nematode EST Project, 1999
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  Nematodes were provided by Dr. Ben-Wen Li and Dr. Gary Weil of
  Washington University, St. Louis, MO (weillab@ingate.wustl.edu).
  Seq primer: -40RP from Gibco.
FEATURES
  source
  1..99
    /organism="Brugia malayi"
    /mol_type="mRNA"
    /db_xref="taxon:6279"
    /dev_stage="L3"
    /lab_host="DH10B"
    /clone_lib="Brugia malayi L3 pAMP1 v2"
    /notes="Vector: pAMPI (Invitrogen); Site_1: NotI; Site_2:
    Sall; The library was constructed by Amy Rush, Claire
    Murphy and Dr. James McCarter at Washington University,
    St. Louis. The cDNA was made by using Dynabead oligo-dT
    priming (Dynal). PCR based library using a modified
    protocol from the SMART PCR cDNA Synthesis Kit from
    Clontech. Directionally cloned into the UDG sites of
    pAMP1. Nematodes were provided by Dr. Ben-Wen Li and Dr.
    Gary Weil of Washington University, St. Louis, MO
    (weillab@ingate.wustl.edu)."
BASE COUNT 31 a 16 c 21 g 31 t
ORIGIN
  Query Match 52.6%; Score 10; DB 14; Length 99;
  Best Local Similarity 52.6%; Pred. No. 2.8e+04;
  Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19
    |||||
Db 97 CCAATAAACTGGCCACG 79

RESULT 12
AZ921866/c
LOCUS
DEFINITION
  AZ921866 99 bp DNA linear GSS 07-JUN-2002
  HRCot4D10 Sorghum bicolor HRCot Sorghum bicolor genomic, genomic
  survey sequence.
ACCESSION
  AZ921866
VERSION
  AZ921866.1 GI:13400225
KEYWORDS
  GSS.
SOURCE
  Sorghum bicolor (sorghum)
  Sorghum bicolor
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
  clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
  1 (bases 1 to 99)
  Peterson,D.G., Schulze,S.R., Sclara,E.B., Lee,S.A., Bowers,J.E.,
  Nagel,A., Jiang,N., Tibbitts,D.C., Weesler,S.R. and Paterson,A.H.
  Integration of Cot analysis, DNA cloning, and high-throughput
  sequencing facilitates genome characterization and gene discovery
  Genome Res. 12 (5), 795-807 (2002)
  21992826
  11997346
  Contact: Peterson DG
  Plant Genome Mapping Laboratory
  University of Georgia
  Room 162, Riverbend Research Bldg., 110 Riverbend Rd., Athens, GA
  30602, USA
  Tel: 706-583-0167
  Fax: 706-583-0160
  Email: dg@arches.uga.edu
  Class: Hydroxyapatite-fractionated DNA.
FEATURES
  Location/Qualifiers
  1..99
    /organism="Sorghum bicolor"
    /mol_type="genomic DNA"
    /cultivar="BTx623"
    /db_xref="taxon:4558"
    /tissue_type="leaves"
    /dev_stage="seedling"
    /clone_lib="Sorghum bicolor HRCot"
    /notes="Vector: pGEN-TA-Easy; A Cot analysis was performed
    for the sorghum genome. Based on the resulting Cot curve,
    hydroxyapatite chromatography was used to isolate
    'highly-repetitive' (HR), 'moderately-repetitive' (MR),
    and 'single/low-copy' (SL) sequence components from
    sheared genomic DNA. The three repetition-based DNA
    components were cloned into E. coli to produce HRCot,
    MRCot, and SLCot genomic libraries. Blotting and
    sequencing data indicates that each library is
    representative of the component from which it was derived.
    Putative ID listings given for sequences are based on
    comparison (blastn) with sequences in the NCBI Nr
    Database. Only the primary match is given (all primary E
    values are < or = 1.00E-5). In no instance does a 'Cot
    clone' contain the complete sequence of its putative Nr
    match."
BASE COUNT 18 a 19 c 29 g 33 t
ORIGIN
  Query Match 52.6%; Score 10; DB 28; Length 99;
  Best Local Similarity 52.6%; Pred. No. 2.8e+04;
  Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19
    |||||
Db 56 CCAATTGCAATCTCCACG 38

RESULT 13
BF086094
LOCUS
DEFINITION
  CM3-GN0052-080900-334-cl1 GN0052 Homo sapiens cDNA, mRNA sequence.
ACCESSION
  BF086094
VERSION
  BF086094.1 GI:10891804
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  ORGANISM

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

#### REFERENCE AUTHORS

1 (bases 1 to 100)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

#### TITLE

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

#### JOURNAL MEDLINE

PROC Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

#### PUBMED COMMENT

20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=et2=CM3-GN0052-080  
900-334-G11at3=2000-09-08&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 5

High quality sequence stop: 100.

Location/Qualifiers

1..100

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="adult"

/clone\_lib="GN0052"

/note="Organ: placenta normal; Vector: puc18; Site 1: SmaI

; Site 2: SmaI; A mini-library was made by cloning

products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer

Research) profiles into the pUC 18 vector. Reverse

transcription of tissue mRNA and cDNA amplification were

performed under low stringency conditions."

35 a 23 c 18 g 24 t

BASE COUNT

ORIGIN

Query Match 52.6%; Score 10; DB 10; Length 100;

Best Local Similarity 52.6%; Pred. No. 2.8e+04;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19

|||||

DB 82 CCAATCTCGAATCCACG 100

|||||

FEATURES

source

RESULT 14

BH583741

LOCUS

BOGXN167R BOGX Brassica oleracea genomic clone BOGXN16, genomic

survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 101)

Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished

JOURNAL

CONTACT: Chris Town

#### TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

#### FEATURES

Location/Qualifiers

1..101

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/strain="T01000DH3"

/db\_xref="taxon:3712"

/clone="BOGXN16"

/clone\_lib="BOGX"

/note="Vector: pHS01; Site 1: BstXI; 2-3 kb sheared

genomic DNA inserted into pHS01 using BstXI linkers"

38 a 23 c 22 g 18 t

BASE COUNT

ORIGIN

Query Match 52.6%; Score 10; DB 28; Length 101;

Best Local Similarity 52.6%; Pred. No. 2.8e+04;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19

|||||

DB 25 CCAATCGCGGCGCCGACG 43

|||||

FEATURES

source

RESULT 15

AA927925

LOCUS

Om23b01.s1 Soares\_NFL\_T\_GBC\_S1 Homo sapiens cDNA clone

IMAGE:1541833 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 553 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 57.

Location/Qualifiers

1..103

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1541833"

/lab\_host="DH10B"

/clone\_lib="Soares\_NFL\_T\_GBC\_S1"

/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with

a modified polylinker; Site 1: Not I; Site 2: Eco RI;

Equal amounts of plasmid DNA from three normalized

libraries (fetal lung NbHL19W, testis NHT, and B-cell

NCI CGAP GCBI) were mixed and ss circles were made in

vitro. Following HAP purification, this DNA was used as

tracer in a subtractive hybridization reaction. The driver

was PCR-amplified cDNAs from pools of 5,000 clones made

from the same 3 libraries. The pools consisted of

I.M.A.G.E. clones 297480-302087, 682632-687239,

726408-728711, and 729096-731399. Subtraction by Bento

Soares and M. Fatima Bonaldo. "

BASE COUNT  
ORIGIN

25 a 23 c 25 g 30 t

Query Match 52.6%; Score 10; DB 9; Length 103;

Best Local Similarity 52.6%; Pred. No. 2.8e+04;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNCCACG 19

Db 55 CCAATTCAATCGGACCACG 37

Search completed: December 4, 2003, 07:47:50  
Job time : 2077 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 08:43:06 ; Search time 2143 Seconds  
(without alignments)  
215.485 Million cell updates/sec

Title: US-09-606-804-1

Perfect score: 19

Sequence: 1 ccaatnnnnnnnccacg 19

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: gb\_est6:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	52.6	43	28	BH857113
C 2	10	52.6	47	29	BX292617
C 3	10	52.6	58	29	AL949322
4	10	52.6	76	9	AA475936

5	10	52.6	79	13	BQ756881
C 6	10	52.6	80	9	AL897168
C 7	10	52.6	80	28	BH251489
8	10	52.6	82	13	BQ099298
C 9	10	52.6	86	28	BH232183
C 10	10	52.6	97	28	BH222106
C 11	10	52.6	99	14	CB338675
C 12	10	52.6	99	28	AZ921866
C 13	10	52.6	100	10	BF086094
14	10	52.6	101	28	BH583741
C 15	10	52.6	103	9	AA927925
C 16	10	52.6	109	28	BH864991
17	10	52.6	110	10	BF094257
18	10	52.6	110	28	AZ720239
19	10	52.6	112	29	BZ662888
C 20	10	52.6	113	28	BH222899
21	10	52.6	114	29	CC458196
22	10	52.6	115	9	AI938545
23	10	52.6	115	10	AW945325
C 24	10	52.6	115	12	BG950375
25	10	52.6	115	13	BW091136
26	10	52.6	115	29	CC029017
27	10	52.6	115	29	AG024826
28	10	52.6	116	29	CC037007
C 29	10	52.6	119	9	AI549504
C 30	10	52.6	120	28	BH230343
C 31	10	52.6	120	28	BH905244
C 32	10	52.6	121	14	CB097190
C 33	10	52.6	121	28	BH222096
34	10	52.6	123	9	AA405167
35	10	52.6	124	28	AQ073902
C 36	10	52.6	124	28	BH812753
C 37	10	52.6	124	28	BH847929
C 38	10	52.6	129	28	AZ919927
39	10	52.6	130	28	BH895687
40	10	52.6	130	29	BZ663067
41	10	52.6	131	29	BZ386507
42	10	52.6	132	10	BF361951
43	10	52.6	132	10	BE936084
44	10	52.6	132	28	BH753521
45	10	52.6	132	28	BH854781

## ALIGNMENTS

RESULT 1

BH857113

LOCUS

DEFINITION

BH857113 43 bp DNA linear GSS 08-JUL-2002  
SALK\_076821.44.95.x Arabidopsis thaliana TDNA insertion lines  
Arabidopsis thaliana genomic clone SALK\_076821.44.95.x, genomic  
survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabisopsis thaliana (thale cress)

Arabisopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu  
This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated exon of At4g29100.  
Class: TDNA tagged.

#### FEATURES

Location/Qualifiers

1..43  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia 0"  
/db\_xref="taxon:3702"

/clone="SALK\_076821.44.95.x"  
/clone\_lib="Arabidopsis thaliana TDNA insertion lines"  
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at [http://signal.salk.edu/tdna\\_protocols.html](http://signal.salk.edu/tdna_protocols.html)"

#### BASE COUNT

11 a 11 c 10 g 11 t

#### ORIGIN

Query Match 52.6%; Score 10; DB 28; Length 43;  
Best Local Similarity 52.6%; Pred. No. 2e+04;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNCCACG 19

Db 18 CCAATGAATACGCCACG 36

#### RESULT 2

BX292617/c

#### LOCUS

DEFINITION BX292617 47 bp DNA linear GSS 07-MAR-2003 Arabidopsis thaliana T-DNA flanking sequence GK-455E08-018778, genomic survey sequence.

ACCESSION BX292617

VERSION BX292617.1

KEYWORDS GI:288991613

SOURCE GSS.

ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1

AUTHORS Strizhov, N., Li, Y., Rosso, M., Viehoever, P., Dekker, K., Siedler, H. and Weisshaar, B.

TITLE A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines

JOURNAL Unpublished

REFERENCE 2

AUTHORS Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weisshaar, B.

TITLE A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics

JOURNAL Unpublished

REFERENCE 3

AUTHORS Li, Y., Strizhov, N., Rosso, M. and Weisshaar, B.

TITLE Direct Submission

JOURNAL Submitted (07-MAR-2003)

COMMENT Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany indicates an insertion within the locus defined by clone f1913. This sequence is recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At1g50730. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

#### FEATURES

Location/Qualifiers

1..47  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia 0"  
/db\_xref="taxon:3702"

/clone="GK-455E08-018778"  
/clone\_lib="Arabidopsis thaliana T-DNA insertion lines"  
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

BASE COUNT 17 a 4 c 10 g 16 t

#### ORIGIN

Query Match 52.6%; Score 10; DB 29; Length 47;

Best Local Similarity 52.6%; Pred. No. 2.1e+04;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNCCACG 19

Db 22 CCATATAAACTCTCCACG 4

#### RESULT 3

AL949322/c

#### LOCUS

DEFINITION AL949322 58 bp DNA linear GSS 24-OCT-2002 Arabidopsis thaliana T-DNA flanking sequence GK-319H03-015860, genomic survey sequence.

ACCESSION AL949322

VERSION AL949322.1

KEYWORDS GI:24405944

SOURCE GSS.

ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1

AUTHORS Strizhov, N., Li, Y., Rosso, M., Viehoever, P., Dekker, K., Siedler, H. and Weisshaar, B.

TITLE A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines

JOURNAL Unpublished

REFERENCE 2

AUTHORS Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weisshaar, B.

TITLE A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics

JOURNAL Unpublished

REFERENCE 3

AUTHORS Rosso, M., Strizhov, N., Li, Y. and Weisshaar, B.

TITLE Direct Submission

JOURNAL Submitted (21-OCT-2002)

COMMENT Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany indicates an insertion close to or within gene At1g50730. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

#### FEATURES

Location/Qualifiers

1..58  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia 0"  
/db\_xref="taxon:3702"  
/clone="GK-319H03-015860"

/clone\_lib="Arabidopsis thaliana T-DNA insertion lines"  
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence

flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

BASE COUNT 21 a 11 c 10 g 16 t

Query Match 52.6%; Score 10; DB 29; Length 58;  
Best Local Similarity 52.6%; Pred. No. 2.3e+04;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19  
|||||

Db 34 CCAATCTGTGTAGCCACG 16  
|||||

RESULT 4  
LOCUS AA475936 76 bp mRNA linear EST 18-JUN-1997  
DEFINITION VM5B10.r1 Soares mammary\_gland NbMMG Mus musculus cDNA clone  
IMAGE:876475 5' similar to TR:E196749 E196749 MRNA; EXPRESSED  
SEQUENCE TAG ;, mRNA sequence.

ACCESSION AA475936  
VERSION AA475936  
KEYWORDS EST  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 76)  
AUTHORS Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T.,  
Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M.,  
Schellenberg K., Steptoe M., Tan F., Underwood K., Moore B.,  
Theising B., Wylie T., Lennon G., Soares B., Wilson R. and  
Waterston R.

TITLE The WashU-HHMI Mouse EST Project  
JOURNAL Unpublished  
COMMENT Contact: Marra M/Mouse EST Project  
WashU-HHMI Mouse EST Project  
Washington University School of MedicineP  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MG1:515955

Trace considered overall poor quality  
Possible reversed clone: similarity on wrong strand  
Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 1.

## FEATURES

Location/Qualifiers  
1..76  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:876475"  
/sex="male"  
/tissue\_type="mammary gland"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/clone\_lib="Soares mammary\_gland NbMMG"  
/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia  
RI; with a modified polylinker; Site 1: Not I; Site 2: Eco  
RI; 1st strand cDNA was primed with a Not I - oligo(dT)  
primer [5,  
TGTTACCAATCTCAAGTCGAGCGCGCGCAATGCTTTTCTTTTCTTTTCTTTT  
T 3']); double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT73 vector.  
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library  
constructed and normalized by Bento Soares and M.Fatima

BASE COUNT 19 a 28 c 16 g 13 t

Query Match 52.6%; Score 10; DB 9; Length 76;  
Best Local Similarity 52.6%; Pred. No. 2.5e+04;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19  
|||||

Db 40 CCAATGGTTTAGCCACG 58  
|||||

## RESULT 5

LOCUS BQ756881

DEFINITION BQ756881

ACCESSION BQ756881

VERSION BQ756881

KEYWORDS EST

SOURCE Hordeum vulgare subsp. vulgare

ORGANISM Hordeum vulgare subsp. vulgare

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

AUTHORS Ramsay L., Machray G., Marshall D.F.M. and Waugh R.

TITLE Development of Barley Transcriptome Resources

JOURNAL Unpublished

COMMENT Contact: Waugh R. Marshall DF  
Genome Dynamics/Computational Biology  
Scottish Crop Research Institute  
Invergowrie, Dundee, DD2 5DA, Scotland, UK  
Tel: 00 44 1382 562731  
Fax: 00 44 1382 562426  
Email: est@scri.sari.ac.uk

FEATURES Location/Qualifiers  
1..79  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="Optic"  
/db\_xref="taxon:112509"  
/clone="EBem09\_SQ005\_E15"  
/tissue\_type="embryo"  
/dev\_stage="1 Day germination"  
/lab\_host="DH10B"  
/clone\_lib="embryo, 1 Day germination, no treatment, cv  
Optic\_EBem09"  
/note="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I;  
Non-normalised library, directionally cloned into pSPORT1.  
Derived from embryos dissected from germinating grains (1  
day) in glasshouse grown barley plants. Developed as part  
of the barley transcriptome resources of BBSRC/SEERAD  
funded cereal IGF (Investigating Gene Function) project."

BASE COUNT 19 a 18 c 17 g 25 t

Query Match 52.6%; Score 10; DB 13; Length 79;  
Best Local Similarity 52.6%; Pred. No. 2.6e+04;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19  
|||||

Db 42 CCAATTTTGTCAACCCACG 60  
|||||

## RESULT 6

LOCUS AL897168/c

DEFINITION AL897168

80 bp mRNA linear EST 16-SEP-2002  
AL897168 XGC-egg Silurana tropicalis cDNA clone TEGG011d06 3', mRNA

```

sequence.
ACCESSION      AL897168
VERSION        AL897168.1  GI:22949622
SOURCE         EST.
ORGANISM       Silurana tropicalis (western clawed frog)
               Silurana tropicalis
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
               Xenopodinae; Silurana.
               1 (bases 1 to 80)
REFERENCE      Taylor,R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.
TITLE          Sanger Xenopus tropicalis EST project 2002
JOURNAL        Unpublished
COMMENT        Contact: Taylor R
               Sanger Centre
               Hinxton, Cambridgeshire, CB10 1SA, UK
               Email: trop@sanger.ac.uk
               Sanger Xenopus tropicalis EST project 2001
               TROPICALIS_SEQUENCE_ID: TEGG011d06.g1kT7
               Sequencing primer: T7
               This sequence is from a Xenopus Gene Collection (XGC) library
               constructed by Aaron M. Zorn.
FEATURES       source
               Location/Qualifiers
               1..80
               /organism="Silurana tropicalis"
               /mol_type="mRNA"
               /db_xref="taxon:8364"
               /clone="TEGG011d06"
               /dev_stage="egg"
               /lab_host="Escherichia coli XL1-blue"
               /clone_lib="XGC-egg"
               /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
               was oligo dt primed from 5'g of poly A+ RNA from egg.
               EcoRI-NotI cut cDNA was then ligated into pCS107 with
               EcoRI at the 5' end and NotI at the 3' end"
BASE COUNT    17 a 12 c 14 g 37 t
ORIGIN
Query Match      52.6%; Score 10; DB 9; Length 80;
Best Local Similarity 52.6%; Pred. No. 2.6e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 71 CCAATCGTAAATGTCACG 53
    |||||

RESULT 7
BH251489/c
LOCUS          BH251489
DEFINITION     SALK_011672 Arabidopsis thaliana TDNA insertion lines Arabidopsis
               thaliana genomic clone SALK_011672, genomic survey sequence.
ACCESSION      BH251489
VERSION        BH251489.1  GI:17138467
KEYWORDS       GSS.
SOURCE         Arabidopsis thaliana (thale cress)
ORGANISM       Arabidopsis thaliana
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
               ; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
               1 (bases 1 to 80)
REFERENCE      Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
               ,C., Jecke,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
               , Zimmerman,J. and Ecker,J.R.
AUTHORS        A Sequence-Indexed Library of Insertion Mutations in the
               Arabidopsis Genome
JOURNAL        Unpublished
COMMENT        Contact: Joseph R. Ecker
               Salk Institute Genomic Analysis Laboratory (SIGnAL)
               The Salk Institute for Biological Studies
               10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
               Tel: 858 453 4100 x1752
               Fax: 858 558 6379

sequence.
ACCESSION      AL897168
VERSION        AL897168.1  GI:22949622
SOURCE         EST.
ORGANISM       Silurana tropicalis (western clawed frog)
               Silurana tropicalis
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
               Xenopodinae; Silurana.
               1 (bases 1 to 80)
REFERENCE      Taylor,R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.
TITLE          Sanger Xenopus tropicalis EST project 2002
JOURNAL        Unpublished
COMMENT        Contact: Taylor R
               Sanger Centre
               Hinxton, Cambridgeshire, CB10 1SA, UK
               Email: trop@sanger.ac.uk
               Sanger Xenopus tropicalis EST project 2001
               TROPICALIS_SEQUENCE_ID: TEGG011d06.g1kT7
               Sequencing primer: T7
               This sequence is from a Xenopus Gene Collection (XGC) library
               constructed by Aaron M. Zorn.
FEATURES       source
               Location/Qualifiers
               1..80
               /organism="Silurana tropicalis"
               /mol_type="mRNA"
               /db_xref="taxon:8364"
               /clone="TEGG011d06"
               /dev_stage="egg"
               /lab_host="Escherichia coli XL1-blue"
               /clone_lib="XGC-egg"
               /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
               was oligo dt primed from 5'g of poly A+ RNA from egg.
               EcoRI-NotI cut cDNA was then ligated into pCS107 with
               EcoRI at the 5' end and NotI at the 3' end"
BASE COUNT    17 a 12 c 14 g 37 t
ORIGIN
Query Match      52.6%; Score 10; DB 9; Length 80;
Best Local Similarity 52.6%; Pred. No. 2.6e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 71 CCAATCGTAAATGTCACG 53
    |||||

RESULT 7
BH251489/c
LOCUS          BH251489
DEFINITION     SALK_011672 Arabidopsis thaliana TDNA insertion lines Arabidopsis
               thaliana genomic clone SALK_011672, genomic survey sequence.
ACCESSION      BH251489
VERSION        BH251489.1  GI:17138467
KEYWORDS       GSS.
SOURCE         Arabidopsis thaliana (thale cress)
ORGANISM       Arabidopsis thaliana
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
               ; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
               1 (bases 1 to 80)
REFERENCE      Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
               ,C., Jecke,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
               , Zimmerman,J. and Ecker,J.R.
AUTHORS        A Sequence-Indexed Library of Insertion Mutations in the
               Arabidopsis Genome
JOURNAL        Unpublished
COMMENT        Contact: Joseph R. Ecker
               Salk Institute Genomic Analysis Laboratory (SIGnAL)
               The Salk Institute for Biological Studies
               10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
               Tel: 858 453 4100 x1752
               Fax: 858 558 6379

Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.
Location/Qualifiers
1..80
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_011672"
/notes="PCR was performed on Arabidopsis thaliana TDNA insertion lines"
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT    18 a 20 c 21 g 19 t 2 others
ORIGIN
Query Match      52.6%; Score 10; DB 28; Length 80;
Best Local Similarity 52.6%; Pred. No. 2.6e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 26 CCAATGATAAATAGCCACG 8
    |||||

RESULT 8
BQ099298
LOCUS          BQ099298
DEFINITION     ph28h02.y2 Ostertagia ostertagi L3 pAMP1 v1 Ostertagia ostertagi
               cDNA 5', mRNA sequence.
ACCESSION      BQ099298
VERSION        BQ099298.1  GI:20132282
KEYWORDS       EST.
SOURCE         Ostertagia ostertagi
ORGANISM       Ostertagia ostertagi
               Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
               Trichostrongyloidea; Haemonchidae; Ostertagiinae; Ostertagia.
               1 (bases 1 to 82)
REFERENCE      McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
               , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
               , Gibbons,M., Ritter,S., Bennett,J., Franklin,C., Teagareishvili,R.
               , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
               , M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
               , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
               Wilson,R.
AUTHORS        The Washington Univ. Nematode EST Project, 1999
               Unpublished
TITLE          Contact: McCarter JP
               The Washington Univ. Nematode EST Project, 1999
JOURNAL        Washington University School of Medicine
               4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
               Tel: 314 286 1800
               Fax: 314 286 1810
               Email: est@wustl.wustl.edu
               The library was constructed by Claire Murphy and Dr. James McCarter
               at Washington University, St. Louis. DNA Sequencing by: Washington
               University Genome Sequencing Center
               High quality sequence stop: 69.
               Location/Qualifiers
               1..82
               /organism="Ostertagia ostertagi"
               /mol_type="mRNA"
               /db_xref="taxon:6317"
               /dev_stage="L3"
               /lab_host="DH10B"
               /clone_lib="Ostertagia ostertagi L3 pAMP1 v1"
               /note="Vector: pAMP1 (Gibco); Site_1: NotI; Site_2: SalI;
               ***WARNING: Subsequent examination of these samples has

```

revealed the presence of an additional Trichostrongyloidea cattle nematode, *Cooperia oncophora*. Sequences in this library may derive from either *Ostertagia* or *Cooperia*. The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. The cDNA was made by using Dynabead oligo-dT priming (Dynal). PCR based library using a modified protocol from the SMART PCR cDNA Synthesis Kit from Clontech. Directionally cloned into the UDG sites of pAMP1. Nematodes were provided by Dr. Louis Gasbarre of the USDA, Beltsville, MD (lgasbar@nri.barc.usda.gov). Third stage exsheathed larvae were collected from 14 day fecal-sphagnum moss cultures of *Ostertagia* eggs. The larvae were recovered by overnight passage on a Baermann apparatus, and then cleaned by passage through a 20 micron nylon mesh. The larvae were then subjected to a treatment with 1.25% chlorox to induce excystation. The larvae were washed with 5 changes of PBS and then pelleted and snap frozen in liquid nitrogen."

```

BASE COUNT      24 a  24 c  12 g  22 t
ORIGIN

Query Match      52.6%; Score 10; DB 13; Length 82;
Best Local Similarity 52.6%; Pred. No. 2.6e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 12 CCAATGGAGAGTGTCACG 30

```

```

RESULT 9
BH232183/c
LOCUS      BH232183      86 bp  DNA  linear  GSS 08-NOV-2001
DEFINITION 1006166D12.Y1 1006 - RescueMu Grid G Zea mays genomic, genomic
survey sequence.
ACCESSION  BH232183
VERSION    BH232183.1 GI:16837119
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays

```

```

REFERENCE
AUTHORS    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
JOURNAL    clade; Panicoideae; Andropogoneae; Zea.
COMMENT    1 (bases 1 to 86)
           Walbot,V.
           Zea mays
           Maize genomic sequences found using engineered RescueMu transposon
           Unpublished
           Contact: Walbot V
           Department of Biological Sciences
           Stanford University
           855 California Ave, Palo Alto, CA 94304, USA
           Tel: 650 723 2227
           Fax: 650 725 8221
           Email: walbot@stanford.edu
           Possible ligation site so sequence was trimmed. Post-ligation
           sequence submitted separately.
           Plate: 1006166 row: 23
           Class: transposon-tagged.
           Location/Qualifiers
             1..86
               /organism="Zea mays"
               /mol_type="genomic DNA"
               /cultivar="mixed background W23/A188/B73"
               /db_xref="taxon:4577"
               /tissue_type="leaf"
               /dev_stage="adult"
               /lab_host="DH10B"
               /clone_lib="1006 - RescueMu Grid G"
               /notes="Organ: leaf; Vector: RescueMu (engineered from
               pBluescript backbone); Site 1: BamHI; Site 2: BglII;
               RescueMu is a 4.9 kb, modified maize Mu transposon
               designed to allow plasmid rescue from total genomic DNA.

```

```

FEATURES
source
1..86
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1006 - RescueMu Grid G"
/notes="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site 1: BamHI; Site 2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.

```

Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu'. Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

```

BASE COUNT      13 a  24 c  37 g  12 t
ORIGIN

```

```

Query Match      52.6%; Score 10; DB 28; Length 86;
Best Local Similarity 52.6%; Pred. No. 2.6e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 61 CCAATCATTTGGACGCCAG 43

```

```

RESULT 10
BH222106/c
LOCUS      BH222106      97 bp  DNA  linear  GSS 08-NOV-2001
DEFINITION 1006105C10.X1 1006 - RescueMu Grid G Zea mays genomic, genomic
survey sequence.
ACCESSION  BH222106
VERSION    BH222106.1 GI:16816606
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays

```

```

REFERENCE
AUTHORS    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
JOURNAL    clade; Panicoideae; Andropogoneae; Zea.
COMMENT    1 (bases 1 to 97)
           Walbot,V.
           Zea mays
           Maize genomic sequences found using engineered RescueMu transposon
           Unpublished
           Contact: Walbot V
           Department of Biological Sciences
           Stanford University
           855 California Ave, Palo Alto, CA 94304, USA
           Tel: 650 723 2227
           Fax: 650 725 8221
           Email: walbot@stanford.edu
           Plate: 1006105 row: 13
           Class: transposon-tagged.
           Location/Qualifiers
             1..97
               /organism="Zea mays"
               /mol_type="genomic DNA"
               /cultivar="mixed background W23/A188/B73"
               /db_xref="taxon:4577"
               /tissue_type="leaf"
               /dev_stage="adult"
               /lab_host="DH10B"
               /clone_lib="1006 - RescueMu Grid G"
               /notes="Organ: leaf; Vector: RescueMu (engineered from
               pBluescript backbone); Site 1: BamHI; Site 2: BglII;
               RescueMu is a 4.9 kb, modified maize Mu transposon
               designed to allow plasmid rescue from total genomic DNA.

```

```

FEATURES
source
1..97
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1006 - RescueMu Grid G"
/notes="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site 1: BamHI; Site 2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu'. Grid G was grown at Stanford in 2000. DNA was
extracted from leaf punches, double digested using BamHI
and BglII, and ligated to form circular plasmids. DH10B
cells were transformed and then screened on LB plates with
ampicillin."
BASE COUNT      11 a  27 c  44 g  15 t
ORIGIN

Query Match      52.6%; Score 10; DB 28; Length 97;
Best Local Similarity 52.6%; Pred. No. 2.8e+04;

```

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNCCACG 19  
 |||||  
 Db 92 CCAATTCGAGCACCACG 74

## RESULT 11

CB38675/c

## LOCUS

DEFINITION CB38675 99 bp mRNA linear EST 13-MAR-2003  
 kb59b04.y1 Brugia malayi L3 pAMP1 v2 Brugia malayi cDNA 5' similar  
 to TR:O16159 O16159 CYSTATIN-TYPE CYSTEINE PROTEINASE INHIBITOR. ;,  
 mRNA sequence.

## ACCESSION

CB38675

## VERSION

CB38675.1

## KEYWORDS

EST.

## SOURCE

Brugia malayi

## ORGANISM

Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 Onchocercidae; Brugia.

## REFERENCE

1 (bases 1 to 99)

## AUTHORS

McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T.,  
 Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y.,  
 Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Teagareishvili, R.,  
 Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe,  
 M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S.,  
 Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and  
 Wilson, R.

## TITLE

The Washington Univ. Nematode EST Project, 1999

## JOURNAL

Unpublished

## COMMENT

Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Nematodes were provided by Dr. Ben-Wen Li and Dr. Gary Weil of

Washington University, St. Louis, MO (weillab@ingate.wustl.edu).

Seq primer: -40RP from Gibco.

## FEATURES

Location/Qualifiers

1..99

/organism="Brugia malayi"

/mol\_type="mRNA"

/db\_xref="taxon:6279"

/dev\_stage="L3"

/lab\_host="DH10B"

/clone\_lib="Brugia malayi L3 pAMP1 v2"

/notes="Vector: pAMP1(Invitrogen); Site.1: NotI; Site.2:  
 SalI; The library was constructed by Amy Rush, Claire  
 Murphy and Dr. James McCarter at Washington University,  
 St. Louis. The cDNA was made by using Dynabead oligo-dT  
 priming (Dynal). PCR based library using a modified  
 protocol from the SMART PCR cDNA Synthesis Kit from  
 Clontech. Directionally cloned into the UDG sites of  
 pAMP1. Nematodes were provided by Dr. Ben-Wen Li and Dr.  
 Gary Weil of Washington University, St. Louis, MO  
 (weillab@ingate.wustl.edu)."

31 a 16 c 21 g 31 c

## BASE COUNT

31 a 16 c 21 g 31 c

## ORIGIN

Query Match

Best Local Similarity 52.6%; Score 10; DB 14; Length 99;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

## Qy

1 CCAATNNNNNNNCCACG 19

|||||

## Db

97 CCAATAAACTTGGCCACG 79

|||||

## RESULT 12

AZ921866/c

## LOCUS

AZ921866

99 bp DNA linear GSS 07-JUN-2002

DEFINITION

BF086094

ACCESSION

BF086094

VERSION

BF086094.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

## DEFINITION

HRCot4D10 Sorghum bicolor HRCot Sorghum bicolor genomic, genomic  
 survey sequence.

## ACCESSION

AZ921866

## VERSION

AZ921866.1

## KEYWORDS

GSS.

## SOURCE

Sorghum bicolor (sorghum)

## ORGANISM

Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Sorghum.

## REFERENCE

1 (bases 1 to 99)

## AUTHORS

Peterson, D.G., Schulze, S.R., Sciara, E.B., Lee, S.A., Bowers, J.E.,  
 Nagel, A., Jiang, N., Tibbitts, D.C., Wessler, S.R. and Paterson, A.H.

## TITLE

Integration of Cot analysis, DNA cloning, and high-throughput  
 sequencing facilitates genome characterization and gene discovery

## JOURNAL

Genome Res. 12 (5), 795-807 (2002)

## MEDLINE

21992826

## PUBMED

11997346

## COMMENT

Contact: Peterson DG

Plant Genome Mapping Laboratory

University of Georgia

Room 162, Riverbend Research Bldg., 110 Riverbend Rd., Athens, GA  
 30602, USA

Tel: 706-583-0167

Fax: 706-583-0160

Email: dgpetarches.uga.edu

Class: Hydroxyapatite-fractionated DNA.

## FEATURES

Location/Qualifiers

1..99

/organism="Sorghum bicolor"

/mol\_type="genomic DNA"

/cultivar="BTx623"

/db\_xref="taxon:4558"

/tissue\_type="leaves"

/dev\_stage="seedling"

/clone\_lib="Sorghum bicolor HRCot"

/notes="Vector: pGEM-TA-Easy; A Cot analysis was performed  
 for the sorghum genome. Based on the resulting Cot curve,  
 hydroxyapatite chromatography was used to isolate  
 'highly-repetitive' (HR), 'moderately-repetitive' (MR),  
 and 'single/low-copy' (SL) sequence components from  
 sheared genomic DNA. The three repetition-based DNA  
 components were cloned into E. coli to produce HRCot,  
 MRCot, and SiCot genomic libraries. Blotting and  
 sequencing data indicates that each library is  
 representative of the component from which it was derived.  
 Putative ID listings given for sequences in the NCBI Nr  
 Database. Only the primary match is given (all primary E  
 values are < or = 1.00E-5). In no instance does a 'Cot  
 clone' contain the complete sequence of its putative Nr  
 match."

18 a 19 c 29 g 33 t

## BASE COUNT

18 a 19 c 29 g 33 t

## ORIGIN

Query Match

Best Local Similarity 52.6%; Score 10; DB 28; Length 99;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

## Qy

1 CCAATNNNNNNNCCACG 19

|||||

## Db

56 CCAATTCGCAATCTCCACG 38

|||||

## RESULT 13

BF086094

## LOCUS

BF086094

## DEFINITION

CM3-GN0052-080900-334-cl1 GN0052 Homo sapiens cDNA, mRNA sequence.

## ACCESSION

BF086094

## VERSION

BF086094.1

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

100 bp mRNA linear EST 19-OCT-2000  
 CM3-GN0052-080900-334-cl1 GN0052 Homo sapiens cDNA, mRNA sequence.

BF086094

DEFINITION

CM3-GN0052-080900-334-cl1 GN0052 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BF086094

VERSION

BF086094.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

#### REFERENCE AUTHORS

1 (bases 1 to 100)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., de Carvalho, A.F., Matukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=kt2-cm3-GN0052-080>)

900-334-cl1kt3=2000-09-08&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 5

High quality sequence stop: 100.

#### FEATURES source

1..100  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="GN0052"  
/note="Organ: placenta normal; Vector: puc18; Site 1: SmaI  
; Site 2: SmaI; A mini-library was made by cloning  
products derived from ORESTES PCR (U.S. Letters Patent  
application No. 196,716 - Ludwig Institute for Cancer  
Research) profiles into the pUC 18 vector. Reverse  
transcription of tissue mRNA and cDNA amplification were  
performed under low stringency conditions."

35 a 23 c 18 g 24 t

#### BASE COUNT ORIGIN

Query Match 52.6%; Score 10; DB 10; Length 100;  
Best Local Similarity 52.6%; Pred. No. 2.8e+04;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19

||||| |||||

DB 82 CCAATTCTCGAATCCACG 100

#### RESULT 14 BH583741

LOCUS BOGXN16TR BOGX Brassica oleracea genomic clone BOGXN16, genomic  
survey sequence.  
DEFINITION

ACCESION BH583741

VERSION BH583741.1 GI:17836198

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 101)

Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished

JOURNAL

CONTACT: Chris Town

Unpublished

CONTACT: Chris Town

Unpublished

CONTACT: Chris Town

Unpublished

CONTACT: Chris Town

#### TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends

Location/Qualifiers

1..101

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/strain="TO1000DH3"

/db\_xref="taxon:3712"

/clone="BOGXN16"

/clone\_lib="BOGX"

/note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared

genomic DNA inserted into PHOS1 using BstXI linkers"

38 a 23 c 22 g 18 t

#### BASE COUNT ORIGIN

Query Match 52.6%; Score 10; DB 28; Length 101;  
Best Local Similarity 52.6%; Pred. No. 2.8e+04;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19

||||| |||||

DB 25 CCAATCCGGCGCCACG 43

||||| |||||

RESULT 15

AA927925/c

LOCUS

DEFINITION

AA927925

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

Insert Length: 553 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 57.

Location/Qualifiers

1..103

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1541833"

/lab\_host="DH10B"

/clone\_lib="Soares NFL T\_GBC\_S1"

/note="Organ: pooled; Vector: pRTD-Pac (Pharmacia) with

a modified polylinker; Site 1: Not I; Site 2: Eco RI;

Equal amounts of plasmid DNA from three normalized

libraries (fetal lung Nhlh19w, testis NHT, and B-cell

NCI CGAP GCBI) were mixed, and as circles were made in

vitro. Following HAP purification, this DNA was used as

tracer in a subtractive hybridization reaction. The driver

was PCR-amplified cDNAs from pools of 5,000 clones made

from the same 3 libraries. The pools consisted of

I.M.A.G.E. clones 297480-302087, 682632-687239,

726408-728711, and 729096-731399. Subtraction by Bento

BASE COUNT 25 a 23 c 25 g 30 t  
ORIGIN Soares and M. Fatima Bonaldo. "

Query Match 52.6%; Score 10; DB 9; Length 103;  
Best Local Similarity 52.6%; Pred. No. 2.8e+04;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNCCACG 19  
|||  
Db 55 CCAATTCATCGGACCACG 37  
|||

Search completed: December 4, 2003, 09:57:25  
Job time : 2150 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 05:16:13 ; Search time 1492 Seconds  
(without alignments)  
520.967 Million cell updates/sec

Title: US-09-606-804-1

Perfect score: 19  
Sequence: 1 ccaatnnnnnnnnccacg 19

Scoring table: IDENTITY NUC  
Gapop 10\*0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: gb\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pin.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	10	52.6	19	6	AR274677 Sequence
2	10	52.6	19	6	AX275341 Sequence
3	10	52.6	19	6	BD016693 Method fo
4	10	52.6	19	6	E28797 Endoplasmic
5	10	52.6	19	6	E28801 Endoplasmic
6	10	52.6	19	6	E28802 Endoplasmic
7	10	52.6	19	6	E28803 Endoplasmic
8	10	52.6	19	6	E28804 Endoplasmic
9	10	52.6	19	6	E28805 Endoplasmic
10	10	52.6	19	6	E28806 Endoplasmic
11	10	52.6	19	6	E28807 Endoplasmic
12	10	52.6	19	6	E28808 Endoplasmic
13	10	52.6	19	6	E28809 Endoplasmic
14	10	52.6	19	6	E44202 Endoplasmic
15	10	52.6	19	6	E44206 Endoplasmic
16	10	52.6	19	6	E44207 Endoplasmic
17	10	52.6	19	6	E44208 Endoplasmic
18	10	52.6	19	6	E44209 Endoplasmic
19	10	52.6	19	6	E44210 Endoplasmic
20	10	52.6	19	6	E44211 Endoplasmic
21	10	52.6	19	6	E44212 Endoplasmic
22	10	52.6	19	6	E44213 Endoplasmic
23	10	52.6	19	6	E44214 Endoplasmic
24	10	52.6	24	6	AX443752 Sequence
25	10	52.6	25	6	AX447733 Sequence
26	10	52.6	29	6	E44234 Endoplasmic
27	10	52.6	40	6	BD180759 Array of
28	10	52.6	99	9	HUMCG3A02
29	10	52.6	105	8	AF479941 Saccharom
30	10	52.6	112	6	AX343726 Sequence
31	10	52.6	118	8	ATH524636 Arabidops
32	10	52.6	122	6	E28800 Endoplasmic
33	10	52.6	126	6	E44205 Endoplasmic
34	10	52.6	125	14	AF271990 Human ade
35	10	52.6	127	6	AX343728 Sequence
36	10	52.6	133	8	ATH523697 Arabidops
37	10	52.6	136	6	AX033516 Arabidops
38	10	52.6	136	6	AX033517 Sequence
39	10	52.6	138	8	AY203001 Arabidops
40	10	52.6	142	6	AX033504 Sequence
41	10	52.6	143	6	AX033505 Sequence
42	10	52.6	143	9	D10339S05 Homo sapie
43	10	52.6	143	9	D10339S05 Sequence
44	10	52.6	145	6	AX072778 Sequence
45	10	52.6	149	6	AX033508 Sequence

#### ALIGNMENTS

RESULT 1  
AR274677  
LOCUS AR274677  
DEFINITION Sequence 1 from patent US 6506600.  
ACCESSION AR274677  
VERSION AR274677.1 GI:29707222  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 19)  
AUTHORS Hermonat,P.L., Mane,M. and Liu,Y.  
TITLE Secreting products from skin by adeno-associated virus (AAV) gene transfer  
JOURNAL Patent: US 6506600-A 1 14-JAN-2003;

```
FEATURES
source
Location/Qualifiers
1..19
/organism="unknown"
BASE COUNT      3 a      5 c      1 g      1 t      9 others
ORIGIN

Query Match      52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.7e+04;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||
Db 1 CCAATNNNNNNNNCCACG 19
    |||

RESULT 2
AX275341
LOCUS      19 bp      DNA      linear      PAT 29-OCT-2001
DEFINITION      Sequence 1 from Patent WO0171018.
ACCESSION      AX275341
VERSION      AX275341.1 GI:16547677
KEYWORDS      .
SOURCE      unidentified
ORGANISM      unidentified
REFERENCE      1
AUTHORS      Hermonat,P.L., Mane,M. and Liu,Y.
TITLE      Secreting products from skin by adeno-associated virus (aav) gene
JOURNAL      Patent: WO 0171018-A 1 27-SEP-2001;
              THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ARKANSAS (US)
FEATURES
source
Location/Qualifiers
1..19
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="glucose response element"
BASE COUNT      3 a      5 c      1 g      1 t      9 others
ORIGIN

Query Match      52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.7e+04;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||
Db 1 CCAATNNNNNNNNCCACG 19
    |||

RESULT 3
BD016693
LOCUS      19 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION      Method for screening ORP150 expression regulatory substance.
ACCESSION      BD016693
VERSION      BD016693.1 GI:22557869
KEYWORDS      JP 2001238699-A/2.
SOURCE      synthetic construct
ORGANISM      synthetic construct
              artificial sequences.
REFERENCE      1 (bases 1 to 19)
AUTHORS      Akazawa,R., Kaneda,S., Yanagagi,H. and Yura,T.
TITLE      Method for screening ORP150 expression regulatory substance
JOURNAL      Patent: JP 2001238699-A 2 04-SEP-2001;
              HSP RESEARCH INST INC
COMMENT      OS Artificial Sequence
              PN JP 2001238699-A/2
              PD 04-SEP-2001
              PF 01-MAR-2000 JP 2000055384
              PI RIEKO AKAZAWA, SUMIKO KANEDA, HIDEKI YANAGI, TAKASHI YURA PC
              C1201/68,A61K45/00,A61P3/10,A61P9/10,A61P17/02,A61P25/ PC
              28,
              PC A61P35/00,A61P37/00,C12N5/10,C12N15/09/(C12N5/10,C12R1:91),
              PC (C12N15/09,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91),
```

```
PC (C12N15/00,C12R1:91)
CC Description of Artificial Sequence:The sequence as shown in
CC SEQ ID NO: 2
CC is ERSE consensus sequence.
CC n is A or C or G or T.
FH Key Location/Qualifiers.
FEATURES
source
Location/Qualifiers
1..19
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT      3 a      5 c      1 g      1 t      9 others
ORIGIN

Query Match      52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.7e+04;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||
Db 1 CCAATNNNNNNNNCCACG 19
    |||

RESULT 4
E28797
LOCUS      19 bp      DNA      linear      PAT 18-JUN-2001
DEFINITION      Endoplasmic reticulum stress-response regulatory element.
ACCESSION      E28797
VERSION      E28797.1 GI:13020851
KEYWORDS      JP 1999243959-A/1.
SOURCE      unidentified
ORGANISM      unidentified
              unclassified.
REFERENCE      1 (bases 1 to 19)
AUTHORS      Hideo,Y., Hideki,Y. and Takashi,Y.
TITLE      Endoplasmic reticulum stress-response regulatory element
JOURNAL      Patent: JP 1999243959-A 1 14-SEP-1999;
              HSP RESEARCH INST INC
COMMENT      OS Unidentified
              PN JP 1999243959-A/1
              PD 14-SEP-1999
              PF 04-MAR-1998 JP 1998052453
              PR
              PI HIDEO YOSHIDA,HIDEKI YANAGI,TAKASHI YURA
              PC C12N15/09,A61K35/74,A61K35/76,A61K38/00,A61K48/00,C12N15/00,
              PC A61K37/02
              CC Strandedness: Double;
              CC Topology: Linear;
              FH Key Location/Qualifiers
              FT enhancer 1..19.
FEATURES
source
Location/Qualifiers
1..19
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT      3 a      5 c      1 g      1 t      9 others
ORIGIN

Query Match      52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.7e+04;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||
Db 1 CCAATNNNNNNNNCCACG 19
    |||

RESULT 5
E28801
LOCUS      19 bp      DNA      linear      PAT 18-JUN-2001
DEFINITION      Endoplasmic reticulum stress-response regulatory element.
ACCESSION      E28801
VERSION      E28801.1 GI:13020855
```

```

KEYWORDS      JP 199243959-A/5.
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 19)
AUTHORS      Hideo,Y., Hideki,Y. and Takashi,Y.
TITLE        Endoplasmic reticulum stress-response regulatory element
JOURNAL      Patent: JP 199243959-A 5 14-SEP-1999;
              HSP RESEARCH INST INC
COMMENT      OS Homo sapiens (human)
              PN JP 199243959-A/5
              PD 14-SEP-1999
              PF 04-MAR-1998 JP 1998052453
              PR
              PI HIDEO YOSHIDA,HIDEKI YANAGI,TAKASHI YURA
              PC C12N15/09,A61K35/74,A61K38/00,A61K48/00,C12N15/00,
              PC A61K37/02
              CC Strandedness: Double;
              CC Topology: Linear;
              FH Key Location/Qualifiers
              FT enhancer 1..19.
FEATURES      source
              1..19
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
BASE COUNT   3 a 9 c 5 g 2 t
ORIGIN
Query Match 52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 52.6%; Pred. No. 2.7e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNCCACG 19
Db 1 CCAATCGGCGCCTCCACG 19

RESULT 6
E28802 LOCUS 19 bp DNA linear PAT 18-JUN-2001
DEFINITION Endoplasmic reticulum stress-response regulatory element.
ACCESSION E28802
VERSION E28802.1 GI:13020856
KEYWORDS JP 199243959-A/6.
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 19)
AUTHORS      Hideo,Y., Hideki,Y. and Takashi,Y.
TITLE        Endoplasmic reticulum stress-response regulatory element
JOURNAL      Patent: JP 199243959-A 6 14-SEP-1999;
              HSP RESEARCH INST INC
COMMENT      OS Mus sp. (mouse)
              PN JP 199243959-A/6
              PD 14-SEP-1999
              PF 04-MAR-1998 JP 1998052453
              PR
              PI HIDEO YOSHIDA,HIDEKI YANAGI,TAKASHI YURA
              PC C12N15/09,A61K35/74,A61K38/00,A61K48/00,C12N15/00,
              PC A61K37/02
              CC Strandedness: Double;
              CC Topology: Linear;
              FH Key Location/Qualifiers
              FT enhancer 1..19.
FEATURES      source
              1..19
              /organism="Mus sp."
              /mol_type="genomic DNA"
              /db_xref="taxon:10095"
BASE COUNT   4 a 8 c 5 g 2 t
ORIGIN
Query Match 52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 52.6%; Pred. No. 2.7e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNCCACG 19
Db 1 CCAATCGGCGCCTCCACG 19

RESULT 8
E28804 LOCUS 19 bp DNA linear PAT 18-JUN-2001
DEFINITION Endoplasmic reticulum stress-response regulatory element.
ACCESSION E28804
VERSION E28804.1 GI:13020858
KEYWORDS JP 199243959-A/8.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 19)
AUTHORS      Hideo,Y., Hideki,Y. and Takashi,Y.
TITLE        Endoplasmic reticulum stress-response regulatory element
JOURNAL      Patent: JP 199243959-A 8 14-SEP-1999;
              HSP RESEARCH INST INC

```

[illegible]

RESULT 10	E28806	E28806	19 bp	DNA	linear	PAT 18-JUN-2001
LOCUS	E28806	E28806	19 bp	DNA	linear	PAT 18-JUN-2001
DEFINITION	Endoplasmic reticulum stress-response regulatory element.	E28806	19 bp	DNA	linear	PAT 18-JUN-2001
ACCESSION	E28806	E28806	19 bp	DNA	linear	PAT 18-JUN-2001
VERSION	E28806.1	GI:13020860	19 bp	DNA	linear	PAT 18-JUN-2001
KEYWORDS	JP 1999243959-A/10.	JP 1999243959-A/10.	19 bp	DNA	linear	PAT 18-JUN-2001
SOURCE	Homo sapiens (human)	Homo sapiens (human)	19 bp	DNA	linear	PAT 18-JUN-2001
ORGANISM	Homo sapiens	Homo sapiens	19 bp	DNA	linear	PAT 18-JUN-2001
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	19 bp	DNA	linear	PAT 18-JUN-2001
AUTHORS	1 (bases 1 to 19)	1 (bases 1 to 19)	19 bp	DNA	linear	PAT 18-JUN-2001
TITLE	Hideo, Y., Hideki, Y. and Takashi, Y.	Hideo, Y., Hideki, Y. and Takashi, Y.	19 bp	DNA	linear	PAT 18-JUN-2001
JOURNAL	Endoplasmic reticulum stress-response regulatory element	Endoplasmic reticulum stress-response regulatory element	19 bp	DNA	linear	PAT 18-JUN-2001
COMMENT	Patent: JP 1999243959-A 10 14-SEP-1999; HSP RESEARCH INST INC	Patent: JP 1999243959-A 10 14-SEP-1999; HSP RESEARCH INST INC	19 bp	DNA	linear	PAT 18-JUN-2001
	OS Homo sapiens (human)	OS Homo sapiens (human)	19 bp	DNA	linear	PAT 18-JUN-2001
	PN JP 1999243959-A/10	PN JP 1999243959-A/10	19 bp	DNA	linear	PAT 18-JUN-2001
	PD 14-SEP-1999	PD 14-SEP-1999	19 bp	DNA	linear	PAT 18-JUN-2001
	PF 04-MAR-1998	PF 04-MAR-1998	19 bp	DNA	linear	PAT 18-JUN-2001
	PR	PR	19 bp	DNA	linear	PAT 18-JUN-2001
	PI HIDEO YOSHIDA, HIDEKI YANAGI, TAKASHI YURA	PI HIDEO YOSHIDA, HIDEKI YANAGI, TAKASHI YURA	19 bp	DNA	linear	PAT 18-JUN-2001
	PC C12N15/09, A61K35/74, A61K35/76, A61K38/00, A61K48/00, C12N15/00, A61K37/02	PC C12N15/09, A61K35/74, A61K35/76, A61K38/00, A61K48/00, C12N15/00, A61K37/02	19 bp	DNA	linear	PAT 18-JUN-2001
	CC Strandedness: Double;	CC Strandedness: Double;	19 bp	DNA	linear	PAT 18-JUN-2001
	CC Topology: Linear;	CC Topology: Linear;	19 bp	DNA	linear	PAT 18-JUN-2001
	FH Key	FH Key	19 bp	DNA	linear	PAT 18-JUN-2001
	FT enhancer	FT enhancer	19 bp	DNA	linear	PAT 18-JUN-2001
FEATURES	Location/Qualifiers	Location/Qualifiers	19 bp	DNA	linear	PAT 18-JUN-2001
source	1..19	1..19	19 bp	DNA	linear	PAT 18-JUN-2001
	/organism="Homo sapiens"	/organism="Homo sapiens"	19 bp	DNA	linear	PAT 18-JUN-2001
	/mol_type="genomic DNA"	/mol_type="genomic DNA"	19 bp	DNA	linear	PAT 18-JUN-2001
	/db_xref="taxon:9606"	/db_xref="taxon:9606"	19 bp	DNA	linear	PAT 18-JUN-2001
BASE COUNT	6 a 6 c 6 g 1 t	6 a 6 c 6 g 1 t	19 bp	DNA	linear	PAT 18-JUN-2001
ORIGIN	1 t	1 t	19 bp	DNA	linear	PAT 18-JUN-2001
	Query Match 52.6%; Score 10; DB 6; Length 19;	Query Match 52.6%; Score 10; DB 6; Length 19;	19 bp	DNA	linear	PAT 18-JUN-2001
	Best Local Similarity 52.6%; Pred. No. 2.7e+04;	Best Local Similarity 52.6%; Pred. No. 2.7e+04;	19 bp	DNA	linear	PAT 18-JUN-2001
	Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;	Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;	19 bp	DNA	linear	PAT 18-JUN-2001
Qy	1 CCAATNNNNNNCCACG 19	1 CCAATNNNNNNCCACG 19	19 bp	DNA	linear	PAT 18-JUN-2001
			19 bp	DNA	linear	PAT 18-JUN-2001
Dd	1 CCAATCGGAGGAGCCACG 19	1 CCAATCGGAGGAGCCACG 19	19 bp	DNA	linear	PAT 18-JUN-2001
			19 bp	DNA	linear	PAT 18-JUN-2001
RESULT 11	E28807	E28807	19 bp	DNA	linear	PAT 18-JUN-2001
LOCUS	E28807	E28807	19 bp	DNA	linear	PAT 18-JUN-2001
DEFINITION	Endoplasmic reticulum stress-response regulatory element.	E28807	19 bp	DNA	linear	PAT 18-JUN-2001
ACCESSION	E28807	E28807	19 bp	DNA	linear	PAT 18-JUN-2001
VERSION	E28807.1	GI:13020861	19 bp	DNA	linear	PAT 18-JUN-2001
KEYWORDS	JP 1999243959-A/11.	JP 1999243959-A/11.	19 bp	DNA	linear	PAT 18-JUN-2001
SOURCE	Gallus sp.	Gallus sp.	19 bp	DNA	linear	PAT 18-JUN-2001
ORGANISM	Gallus sp.	Gallus sp.	19 bp	DNA	linear	PAT 18-JUN-2001
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.	19 bp	DNA	linear	PAT 18-JUN-2001
AUTHORS	1 (bases 1 to 19)	1 (bases 1 to 19)	19 bp	DNA	linear	PAT 18-JUN-2001
TITLE	Hideo, Y., Hideki, Y. and Takashi, Y.	Hideo, Y., Hideki, Y. and Takashi, Y.	19 bp	DNA	linear	PAT 18-JUN-2001
JOURNAL	Endoplasmic reticulum stress-response regulatory element	Endoplasmic reticulum stress-response regulatory element	19 bp	DNA	linear	

```

CC Topology: Linear;
FH Key Location/Qualifiers
FT enhancer 1..19.

FEATURES
    source
        Location/Qualifiers
            1..19
            /organism="Gallus sp."
            /mol_type="genomic DNA"
            /db_xref="taxon:9036"
BASE COUNT      4 a      9 c      5 g      1 t
ORIGIN
Query Match      52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 52.6%; Pred. No. 2.7e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 1 CCAATCGACGCCGCCACG 19
    |||||

RESULT 12
LOCUS      E28808      19 bp DNA linear PAT 18-JUN-2001
DEFINITION Endoplasmic reticulum stress-response regulatory element.
ACCESSION E28808
VERSION JP 1999243959-A/12.
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 19)
Hideo, Y., Hideo, Y. and Takaishi, Y.
Endoplasmic reticulum stress-response regulatory element
Patent: JP 1999243959-A 12 14-SEP-1999;
HSP RESEARCH INST INC
OS Homo sapiens (human)
PN JP 1999243959-A/12
PD 14-SEP-1999
PF 04-MAR-1998 JP 1998052453
PR HIDEO YOSHIDA, HIDEKI YANAGI, TAKASHI YURA
PC C12N15/09, A61K35/74, A61K38/00, A61K48/00, C12N15/00,
A61K37/02
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT enhancer 1..19.

FEATURES
    source
        Location/Qualifiers
            1..19
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT      5 a      6 c      5 g      3 t
ORIGIN
Query Match      52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 52.6%; Pred. No. 2.7e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 1 CCAATGATGTCGACACG 19
    |||||

RESULT 13
LOCUS      E28809      19 bp DNA linear PAT 18-JUN-2001
DEFINITION Endoplasmic reticulum stress-response regulatory element.
ACCESSION E28809
VERSION JP 1999243959-A/13.
KEYWORDS Mus sp.
SOURCE Mus sp.

```

```

ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)
Hideo, Y., Hideo, Y. and Takaishi, Y.
Endoplasmic reticulum stress-response regulatory element
Patent: JP 1999243959-A 13 14-SEP-1999;
HSP RESEARCH INST INC
OS Mus sp. (mouse)
PN JP 1999243959-A/13
PD 14-SEP-1999
PF 04-MAR-1998 JP 1998052453
PR HIDEO YOSHIDA, HIDEKI YANAGI, TAKASHI YURA
PC C12N15/09, A61K35/74, A61K38/00, A61K48/00, C12N15/00,
A61K37/02
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT enhancer 1..19.

FEATURES
    source
        Location/Qualifiers
            1..19
            /organism="Mus sp."
            /mol_type="genomic DNA"
            /db_xref="taxon:10095"
BASE COUNT      5 a      6 c      6 g      2 t
ORIGIN
Query Match      52.6%; Score 10; DB 6; Length 19;
Best Local Similarity 52.6%; Pred. No. 2.7e+04;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 1 CCAATGAGGTCGACACG 19
    |||||

RESULT 14
LOCUS      E44202      19 bp DNA linear PAT 31-JAN-2002
DEFINITION Endoplasmic reticulum stress transcription factor.
ACCESSION E44202
VERSION E44202.1 GI:18633455
KEYWORDS JP 2001054391-A/1.
SOURCE synthetic construct
ORGANISM artificial sequences.
1 (bases 1 to 19)
Haji, K., Yoshida, H., Mori, K., Yanagi, H. and Yura, T.
Endoplasmic reticulum stress transcription factor
Patent: JP 2001054391-A 1 27-FEB-2001;
HSP RESEARCH INST INC
OS Artificial Sequence
PN JP 2001054391-A/1
PD 27-FEB-2001
PF 11-NOV-1999 JP 1999321743
PR KYOSUKE HAJI, HIDEO YOSHIDA, KAZUTOSHI MORI, HIDEKI YANAGI, PI
TAKASHI YURA
PC C12N15/09, C12P21/02, (C12N15/09, C12R1:91), C12N15/00, (C12N15/00, PC
C12R1:91)
CC
FH Key Location/Qualifiers
FT source 1..19
    /organism="Artificial Sequence".

FEATURES
    source
        Location/Qualifiers
            1..19
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
BASE COUNT      3 a      5 c      1 g      1 t      9 others
ORIGIN

```

Query Match 52.6%; Score 10; DB 6; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.7e+04;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19  
|||||  
Db 1 CCAATNNNNNNCCACG 19

RESULT 15

E44206 19 bp DNA linear PAT 31-JAN-2002  
LOCUS Endoplasmic reticulum stress transcription factor.  
DEFINITION  
E44206  
E44206.1 GI:18633459  
VERSION JP 2001054391-A/5.  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 19)  
Haji,K., Yoshida,H., Mori,K., Yanagi,H. and Yura,T.  
Endoplasmic reticulum stress transcription factor  
Patent: JP 2001054391-A 5 27-FEB-2001;  
HSP RESEARCH INST INC

OS Homo sapiens (human)  
PN JP 2001054391-A/5  
PD 27-FEB-2001  
PF 11-NOV-1999 JP 1999321743  
PR

PI KYOSUKE HAJI,HIDEO YOSHIDA,KAZUTOSHI MORI,HIDEKI YANAGI, PI  
TAKASHI YURA  
PC  
C12N15/09,C12P21/02//(C12N15/09,C12R1:91),C12N15/00,(C12N15/00, PC  
C12R1:91)

CC  
FH Key Location/Qualifiers  
FT source 1..19 /organism='Homo sapiens (human)'.  
FEATURES  
source Location/Qualifiers  
1..19  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

BASE COUNT 3 a 9 c 5 g 2 t  
ORIGIN

Query Match 52.6%; Score 10; DB 6; Length 19;  
Best Local Similarity 52.6%; Pred. No. 2.7e+04;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19  
|||||  
Db 1 CCAATCGCGGCTCCACG 19

Search completed: December 4, 2003, 07:13:02  
Job time : 1495 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 04:27:16 ; Search time 256 seconds  
(without alignments)  
200.349 Million cell updates/sec

Title: US-09-606-804-1

Perfect score: 19

Sequence: 1 ccaatnnnnnnnnccacg 19

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03:\*

1: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	52.6	19	20	AAZ25631
2	10	52.6	19	20	AAZ25632
3	10	52.6	19	20	AAZ25633
4	10	52.6	19	20	AAZ25634
5	10	52.6	19	20	AAZ25635
6	10	52.6	19	20	AAZ25636
7	10	52.6	19	20	AAZ25637
8	10	52.6	19	20	AAZ25638

9	10	52.6	19	20	AAZ25639	Endoplasmic reticu
10	10	52.6	19	20	AAZ25640	Endoplasmic reticu
11	10	52.6	19	21	AAA28570	GRP78 promoter ERS
12	10	52.6	19	21	AAA28571	GRP78 promoter ERS
13	10	52.6	19	21	AAA28572	GRP78 promoter ERS
14	10	52.6	19	21	AAA28573	GRP94 promoter ERS
15	10	52.6	19	21	AAA28574	GRP94 promoter ERS
16	10	52.6	19	21	AAA28575	GRP94 promoter ERS
17	10	52.6	19	21	AAA28576	GRP94 promoter ERS
18	10	52.6	19	21	AAA28577	Calreticulin promo
19	10	52.6	19	21	AAA28578	Calreticulin promo
20	10	52.6	19	21	AAA28579	ERSE consensus se
21	10	52.6	19	22	AAI70001	ERSE consensus se
22	10	52.6	19	22	AAI70002	Endoplasmic respon
23	10	52.6	24	21	AAA28596	GRP78 promoter ERS
24	10	52.6	24	21	ABQ00200	Oligonucleotide ad
25	10	52.6	24	24	ABQ04398	Oligonucleotide ad
26	10	52.6	24	24	ABQ04439	Oligonucleotide ad
27	10	52.6	24	24	ABQ10685	Oligonucleotide ad
28	10	52.6	24	24	ABQ10726	Oligonucleotide ad
29	10	52.6	25	19	AAV45529	Helicobacter pylor
30	10	52.6	25	24	ABQ12221	Oligonucleotide ad
31	10	52.6	25	24	ABQ12262	Oligonucleotide ad
32	10	52.6	28	21	AAA28597	GRP94 promoter ERS
33	10	52.6	28	21	AAA28599	Calreticulin promo
34	10	52.6	29	21	AAA28602	Tandem repeat comp
35	10	52.6	30	21	AAA28598	GRP94 promoter ERS
36	10	52.6	33	21	AAA28569	GRP78 promoter tan
37	10	52.6	50	24	ABZ03658	Human leukocyte ge
38	10	52.6	112	24	ABL41923	Nucleotide sequenc
39	10	52.6	122	20	AAZ25660	Human GRP78 promot
40	10	52.6	122	21	AAA28586	5' flanking region
41	10	52.6	123	24	ABL60984	Human ovarian canc
42	10	52.6	126	21	AAV471906	IAK fragment NS fr
43	10	52.6	127	24	ABL41924	Nucleotide sequenc
44	10	52.6	143	21	AAA71907	IAK fragment NS fr
45	10	52.6	145	22	AAAF67488	Novel human polynu

## ALIGNMENTS

RESULT 1

AAZ25631

ID AAZ25631 standard; DNA; 19 BP.

XX AAZ25631;

AC AC

XX 23-DEC-1999 (first entry)

DT DT

XX Endoplasmic reticulum stress competence control element SEQ ID NO:1.

DE Endoplasmic reticulum; ER; stress competence; control element;

XX Endoplasmic reticulum; ER; stress competence; control element;

KW inhibition; growth; apoptosis; cancer; autoimmune disease;

KW cystic fibrosis; ds.

XX Homo sapiens.

OS Homo sapiens.

XX JP11243959-A.

PN JP11243959-A.

XX 14-SEP-1999.

XX 04-MAR-1998; 98JP-0052453.

XX 04-MAR-1998; 98JP-0052453.

XX (HSPK-) HSP KENKYUSHO KK.

PA WPI; 1999-603708/52.

XX New control element for stress competence of endoplasmic reticulum -

XX useful for inhibition of growth and induction of apoptosis in cancer

PT cells

XX PS Claim 1; Page 10; 25pp; Japanese.

XX CC The present invention specifically claims an element shown by: (A) a  
CC 19 bp base sequence, CCAATNNNNNNNNCCACG (ERSE); or (B) a modified base  
CC sequence having replaced 1-3 bases with the other base(s), which induces  
CC transcription with stress on endoplasmic reticulum used for stress  
CC competence of endoplasmic reticulum. Also described are: (1) a DNA having  
CC transcription inducing activity with stress on endoplasmic reticulum  
CC containing the above mentioned element, optionally further containing a  
CC promoter DNA; and (2) a vector containing the element optionally with the  
CC DNA. The element can be used for the inhibition of growth and induction  
CC of apoptosis of cancer cells, and improvement of symptoms of autoimmune  
CC diseases and cystic fibrosis by inhibition of autoantibody formation.

XX SQ Sequence 19 BP; 3 A; 5 C; 1 G; 1 T; 9 other;

Query Match 52.6%; Score 10; DB 20; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
Db 1 CCAATNNNNNNNNCCACG 19  
|||||

RESULT 2  
AAZ25632  
ID AAZ25632 standard; DNA; 19 BP.  
XX AC AAZ25632;  
XX DT 23-DEC-1999 (first entry)  
XX DE Endoplasmic reticulum stress competence control element SEQ ID NO:5.  
XX KW Endoplasmic reticulum; ER; stress competence; control element;  
XX inhibition; growth; apoptosis; cancer; autoimmune disease;  
XX cystic fibrosis; ds.  
XX OS Homo sapiens.  
XX PN JP11243959-A.  
XX PD 14-SEP-1999.  
XX PF 04-MAR-1998; 98JP-0052453.  
XX PR 04-MAR-1998; 98JP-0052453.  
XX PA (HSPK-) HSP KENKYUSHO KK.  
XX DR WPI; 1999-603708/52.  
XX PT New control element for stress competence of endoplasmic reticulum -  
PT useful for inhibition of growth and induction of apoptosis in cancer  
PT cells  
XX PS Example 1; Fig 3; 25pp; Japanese.

XX CC The present invention specifically claims an element shown by: (A) a  
CC 19 bp base sequence, CCAATNNNNNNNNCCACG (ERSE); or (B) a modified base  
CC sequence having replaced 1-3 bases with the other base(s), which induces  
CC transcription with stress on endoplasmic reticulum used for stress  
CC competence of endoplasmic reticulum. Also described are: (1) a DNA  
CC having transcription inducing activity with stress on endoplasmic  
CC reticulum containing the above mentioned element, optionally further  
CC containing a promoter DNA; and (2) a vector containing the element  
CC optionally with the DNA. The element can be used for the inhibition of  
CC growth and induction of apoptosis of cancer cells, and improvement of  
CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of  
CC autoantibody formation. AAZ25632 to AAZ25657 represent elements used in  
CC an example from the present invention.

XX SQ Sequence 19 BP; 3 A; 9 C; 5 G; 2 T; 0 other;

Query Match 52.6%; Score 10; DB 20; Length 19;  
Best Local Similarity 52.6%; Pred. No. 1.8e+03;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
Db 1 CCAATCGGCGCTCCACG 19  
|||||

RESULT 3  
AAZ25633  
ID AAZ25633 standard; DNA; 19 BP.  
XX AC AAZ25633;  
XX DT 23-DEC-1999 (first entry)  
XX DE Endoplasmic reticulum stress competence control element SEQ ID NO:6.  
XX KW Endoplasmic reticulum; ER; stress competence; control element;  
XX inhibition; growth; apoptosis; cancer; autoimmune disease;  
XX cystic fibrosis; ds.  
XX OS Mus sp.  
XX PN JP11243959-A.  
XX PD 14-SEP-1999.  
XX PF 04-MAR-1998; 98JP-0052453.  
XX PR 04-MAR-1998; 98JP-0052453.  
XX PA (HSPK-) HSP KENKYUSHO KK.  
XX DR WPI; 1999-603708/52.  
XX PT New control element for stress competence of endoplasmic reticulum -  
PT useful for inhibition of growth and induction of apoptosis in cancer  
PT cells  
XX PS Example 1; Fig 3; 25pp; Japanese.

XX CC The present invention specifically claims an element shown by: (A) a  
CC 19 bp base sequence, CCAATNNNNNNNNCCACG (ERSE); or (B) a modified base  
CC sequence having replaced 1-3 bases with the other base(s), which induces  
CC transcription with stress on endoplasmic reticulum used for stress  
CC competence of endoplasmic reticulum. Also described are: (1) a DNA  
CC having transcription inducing activity with stress on endoplasmic  
CC reticulum containing the above mentioned element, optionally further  
CC containing a promoter DNA; and (2) a vector containing the element  
CC optionally with the DNA. The element can be used for the inhibition of  
CC growth and induction of apoptosis of cancer cells, and improvement of  
CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of  
CC autoantibody formation. AAZ25632 to AAZ25657 represent elements used in  
CC an example from the present invention.

XX SQ Sequence 19 BP; 4 A; 8 C; 5 G; 2 T; 0 other;

Query Match 52.6%; Score 10; DB 20; Length 19;  
Best Local Similarity 52.6%; Pred. No. 1.8e+03;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
Db 1 CCAATCGGCGCTCCACG 19  
|||||

RESULT 4  
AAZ25634



```

ID AA225634 standard; DNA; 19 BP.
XX AC
XX AA225634;
XX DT
XX 23-DEC-1999 (first entry)
XX DE
XX Endoplasmic reticulum stress competence control element SEQ ID NO:7.
XX KW
XX Endoplasmic reticulum; ER; stress competence; control element;
XX KW inhibition; growth; apoptosis; cancer; autoimmune disease;
XX KW cystic fibrosis; ds.
XX OS
XX Rattus sp.
XX XX
XX JPI1243959-A.
XX PN
XX 14-SEP-1999.
XX PD
XX 04-MAR-1998; 98JP-0052453.
XX PF
XX 04-MAR-1998; 98JP-0052453.
XX PR
XX (HSPK-) HSP KENKYUSHO KK.
XX PA
XX WPI; 1999-603708/52.
XX DR
XX New control element for stress competence of endoplasmic reticulum -
XX PT useful for inhibition of growth and induction of apoptosis in cancer
XX PT cells
XX PS
XX Example 1; Fig 3; 25pp; Japanese.
XX CC
XX The present invention specifically claims an element shown by: (A) a
XX CC 19 bp base sequence, CCAATNNNNN NNNCCAG (ERSE); or (B) a modified base
XX CC sequence having replaced 1-3 bases with the other base(s), which induces
XX CC transcription with stress on endoplasmic reticulum used for stress
XX CC competence of endoplasmic reticulum. Also described are: (1) a DNA
XX CC having transcription inducing activity with stress on endoplasmic
XX CC reticulum containing the above mentioned element, optionally further
XX CC containing a promoter DNA; and (2) a vector containing the element
XX CC optionally with the DNA. The element can be used for the inhibition of
XX CC growth and induction of apoptosis of cancer cells, and improvement of
XX CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
XX CC autoantibody formation. AA225632 to AA225657 represent elements used in
XX CC an example from the present invention.
XX SQ
XX Sequence 19 BP; 4 A; 8 C; 5 G; 2 T; 0 other;

Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNN NNNCCAG 19
Db 1 CCAATCGAGGCTCCAG 19

RESULT 5
AAZ25635
ID AA225635 standard; DNA; 19 BP.
XX AC
XX AA225635;
XX DT
XX 23-DEC-1999 (first entry)
XX DE
XX Endoplasmic reticulum stress competence control element SEQ ID NO:8.
XX KW
XX Endoplasmic reticulum; ER; stress competence; control element;
XX KW inhibition; growth; apoptosis; cancer; autoimmune disease;
XX KW cystic fibrosis; ds.
XX OS
XX Homo sapiens.
XX PT

```

```

PN JPI1243959-A.
XX PD
XX 14-SEP-1999.
XX PF
XX 04-MAR-1998; 98JP-0052453.
XX PR
XX 04-MAR-1998; 98JP-0052453.
XX PA
XX (HSPK-) HSP KENKYUSHO KK.
XX XX
XX WPI; 1999-603708/52.
XX XX
XX New control element for stress competence of endoplasmic reticulum -
XX PT useful for inhibition of growth and induction of apoptosis in cancer
XX PT cells
XX PS
XX Example 1; Fig 3; 25pp; Japanese.
XX CC
XX The present invention specifically claims an element shown by: (A) a
XX CC 19 bp base sequence, CCAATNNNNN NNNCCAG (ERSE); or (B) a modified base
XX CC sequence having replaced 1-3 bases with the other base(s), which induces
XX CC transcription with stress on endoplasmic reticulum used for stress
XX CC competence of endoplasmic reticulum. Also described are: (1) a DNA
XX CC having transcription inducing activity with stress on endoplasmic
XX CC reticulum containing the above mentioned element, optionally further
XX CC containing a promoter DNA; and (2) a vector containing the element
XX CC optionally with the DNA. The element can be used for the inhibition of
XX CC growth and induction of apoptosis of cancer cells, and improvement of
XX CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
XX CC autoantibody formation. AA225632 to AA225657 represent elements used in
XX CC an example from the present invention.
XX SQ
XX Sequence 19 BP; 4 A; 10 C; 4 G; 1 T; 0 other;

Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNN NNNCCAG 19
Db 1 CCAATCGCGCCGACCAG 19

RESULT 6
AAZ25636
ID AA225636 standard; DNA; 19 BP.
XX AC
XX AA225636;
XX DT
XX 23-DEC-1999 (first entry)
XX DE
XX Endoplasmic reticulum stress competence control element SEQ ID NO:9.
XX KW
XX Endoplasmic reticulum; ER; stress competence; control element;
XX KW inhibition; growth; apoptosis; cancer; autoimmune disease;
XX KW cystic fibrosis; ds.
XX OS
XX Gallus sp.
XX XX
XX JPI1243959-A.
XX PN
XX 14-SEP-1999.
XX PD
XX 04-MAR-1998; 98JP-0052453.
XX PF
XX 04-MAR-1998; 98JP-0052453.
XX PR
XX (HSPK-) HSP KENKYUSHO KK.
XX PA
XX WPI; 1999-603708/52.
XX DR
XX New control element for stress competence of endoplasmic reticulum -
XX PT useful for inhibition of growth and induction of apoptosis in cancer
XX PT cells

```

```
PT cells
XX
PS Example 1; Fig 3; 25pp; Japanese.
XX
CC The present invention specifically claims an element shown by: (A) a
CC 19 bp base sequence, CCAATNNNN NNNCCACG (ERSE); or (B) a modified base
CC sequence having replaced 1-3 bases with the other base(s), which induces
CC transcription with stress on endoplasmic reticulum used for stress
CC competence of endoplasmic reticulum. Also described are: (1) a DNA
CC having transcription inducing activity with stress on endoplasmic
CC reticulum containing the above mentioned element, optionally further
CC containing a promoter DNA; and (2) a vector containing the element
CC optionally with the DNA. The element can be used for the inhibition of
CC growth and induction of apoptosis of cancer cells, and improvement of
CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
CC autoantibody formation. AA225632 to AA225657 represent elements used in
CC an example from the present invention.
XX
SQ Sequence 19 BP; 5 A; 7 C; 6 G; 1 T; 0 other;
Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 CCAATNNNNNNNNCCACG 19
Db 1 CCAATGGGAGCGCACACG 19
RESULT 7
AA225637
ID AA225637 standard; DNA; 19 BP.
XX
AC AA225637;
XX
DT 23-DEC-1999 (first entry)
XX
DE Endoplasmic reticulum stress competence control element SEQ ID NO:10.
XX
KW Endoplasmic reticulum; ER; stress competence; control element;
KW inhibition; growth; apoptosis; cancer; autoimmune disease;
KW cystic fibrosis; ds.
XX
OS Homo sapiens.
XX
PN JP11243959-A.
XX
PD 14-SEP-1999.
XX
PF 04-MAR-1998; 98JP-0052453.
XX
PR 04-MAR-1998; 98JP-0052453.
XX
PA (HSPK-) HSP KENKYUSHO KK.
XX
DR WPI; 1999-603708/52.
XX
PT New control element for stress competence of endoplasmic reticulum -
PT useful for inhibition of growth and induction of apoptosis in cancer
PT cells
XX
PS Example 1; Fig 3; 25pp; Japanese.
XX
CC The present invention specifically claims an element shown by: (A) a
CC 19 bp base sequence, CCAATNNNN NNNCCACG (ERSE); or (B) a modified base
CC sequence having replaced 1-3 bases with the other base(s), which induces
CC transcription with stress on endoplasmic reticulum used for stress
CC competence of endoplasmic reticulum. Also described are: (1) a DNA
CC having transcription inducing activity with stress on endoplasmic
CC reticulum containing the above mentioned element, optionally further
CC containing a promoter DNA; and (2) a vector containing the element
CC optionally with the DNA. The element can be used for the inhibition of
CC growth and induction of apoptosis of cancer cells, and improvement of
CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
CC autoantibody formation. AA225632 to AA225657 represent elements used in
CC an example from the present invention.
XX
SQ Sequence 19 BP; 6 A; 6 C; 6 G; 1 T; 0 other;
Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 CCAATNNNNNNNNCCACG 19
Db 1 CCAATGGGAGCGCACACG 19
RESULT 8
AA225638
ID AA225638 standard; DNA; 19 BP.
XX
AC AA225638;
XX
DT 23-DEC-1999 (first entry)
XX
DE Endoplasmic reticulum stress competence control element SEQ ID NO:11.
XX
KW Endoplasmic reticulum; ER; stress competence; control element;
KW inhibition; growth; apoptosis; cancer; autoimmune disease;
KW cystic fibrosis; ds.
XX
OS Gallus sp.
XX
PN JP11243959-A.
XX
PD 14-SEP-1999.
XX
PF 04-MAR-1998; 98JP-0052453.
XX
PR 04-MAR-1998; 98JP-0052453.
XX
PA (HSPK-) HSP KENKYUSHO KK.
XX
DR WPI; 1999-603708/52.
XX
PT New control element for stress competence of endoplasmic reticulum -
PT useful for inhibition of growth and induction of apoptosis in cancer
PT cells
XX
PS Example 1; Fig 3; 25pp; Japanese.
XX
CC The present invention specifically claims an element shown by: (A) a
CC 19 bp base sequence, CCAATNNNN NNNCCACG (ERSE); or (B) a modified base
CC sequence having replaced 1-3 bases with the other base(s), which induces
CC transcription with stress on endoplasmic reticulum used for stress
CC competence of endoplasmic reticulum. Also described are: (1) a DNA
CC having transcription inducing activity with stress on endoplasmic
CC reticulum containing the above mentioned element, optionally further
CC containing a promoter DNA; and (2) a vector containing the element
CC optionally with the DNA. The element can be used for the inhibition of
CC growth and induction of apoptosis of cancer cells, and improvement of
CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
CC autoantibody formation. AA225632 to AA225657 represent elements used in
CC an example from the present invention.
XX
SQ Sequence 19 BP; 4 A; 9 C; 5 G; 1 T; 0 other;
Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 CCAATNNNNNNNNCCACG 19
Db 1 CCAATGGGAGCGCGCACG 19
```

```

RESULT 9
AAZ25639
ID AAZ25639 standard; DNA; 19 BP.
XX
AC AAZ25639;
XX
DT 23-DEC-1999 (first entry)
XX
DE Endoplasmic reticulum stress competence control element SEQ ID NO:12.
XX
KW Endoplasmic reticulum; ER; stress competence; control element;
KW inhibition; growth; apoptosis; cancer; autoimmune disease;
KW cystic fibrosis; ds.
XX
OS Homo sapiens.
XX
PN JP11243959-A.
XX
PD 14-SEP-1999.
XX
PF 04-MAR-1998; 98JP-0052453.
XX
PR 04-MAR-1998; 98JP-0052453.
XX
PA (HSPK-) HSP KENKYUSHO KK.
XX
DR WPI; 1999-603708/52.
XX
PT New control element for stress competence of endoplasmic reticulum -
PT useful for inhibition of growth and induction of apoptosis in cancer
PT cells
XX
PS Example 1; Fig 3; 25pp; Japanese.
XX
CC The present invention specifically claims an element shown by: (A) a
CC 19 bp base sequence, CCAATNNNN NNNCCACG (ERSE); or (B) a modified base
CC sequence having replaced 1-3 bases with the other base(s), which induces
CC transcripion with stress on endoplasmic reticulum used for stress
CC competence of endoplasmic reticulum. Also described are: (1) a DNA
CC having transcription inducing activity with stress on endoplasmic
CC reticulum containing the above mentioned element, optionally further
CC containing a promoter DNA; and (2) a vector containing the element
CC optionally with the DNA. The element can be used for the inhibition of
CC growth and induction of apoptosis of cancer cells, and improvement of
CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
CC autoantibody formation. AAZ25632 to AAZ25657 represent elements used in
CC an example from the present invention.
XX
SQ Sequence 19 BP; 5 A; 6 C; 5 G; 3 T; 0 other;

Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||||
DB 1 CCAATGATGTCGACCG 19
    |||||

RESULT 10
AAZ25640
ID AAZ25640 standard; DNA; 19 BP.
XX
AC AAZ25640;
XX
DT 23-DEC-1999 (first entry)
XX
DE Endoplasmic reticulum stress competence control element SEQ ID NO:13.
XX
KW Endoplasmic reticulum; ER; stress competence; control element;
KW inhibition; growth; apoptosis; cancer; autoimmune disease;
KW cystic fibrosis; ds.
XX
OS Homo sapiens.
XX
PN JP11243959-A.
XX
PD 14-SEP-1999.
XX
PF 04-MAR-1998; 98JP-0052453.
XX
PR 04-MAR-1998; 98JP-0052453.
XX
PA (HSPK-) HSP KENKYUSHO KK.
XX
DR WPI; 1999-603708/52.
XX
PT New control element for stress competence of endoplasmic reticulum -
PT useful for inhibition of growth and induction of apoptosis in cancer
PT cells
XX
PS Example 1; Fig 3; 25pp; Japanese.
XX
CC The present invention specifically claims an element shown by: (A) a
CC 19 bp base sequence, CCAATNNNN NNNCCACG (ERSE); or (B) a modified base
CC sequence having replaced 1-3 bases with the other base(s), which induces
CC transcripion with stress on endoplasmic reticulum used for stress
CC competence of endoplasmic reticulum. Also described are: (1) a DNA
CC having transcription inducing activity with stress on endoplasmic
CC reticulum containing the above mentioned element, optionally further
CC containing a promoter DNA; and (2) a vector containing the element
CC optionally with the DNA. The element can be used for the inhibition of
CC growth and induction of apoptosis of cancer cells, and improvement of
CC symptoms of autoimmune diseases and cystic fibrosis by inhibition of
CC autoantibody formation. AAZ25632 to AAZ25657 represent elements used in
CC an example from the present invention.
XX
SQ Sequence 19 BP; 5 A; 6 C; 5 G; 3 T; 0 other;

Query Match 52.6%; Score 10; DB 20; Length 19;
Best Local Similarity 52.6%; Pred. No. 1.8e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
    |||||
DB 1 CCAATGATGTCGACCG 19
    |||||

RESULT 11
AAZ28570
ID AAA28570 standard; DNA; 19 BP.
XX
AC AAA28570;
XX
DT 29-AUG-2000 (first entry)
XX
DE GRP78 promoter ERSE1-like sequence.
XX
KW Endoplasmic reticulum; stress; ER; transcription factor;
KW transcripion; regulatory element; ERSE; bZIP; chaperone; treatment;
KW prophylaxis; cancer; arteriosclerosis; ischaemia; wound healing;
KW cystic fibrosis; ulcer; gene therapy; recombinant gene; human;
KW gene expression; GRP; glucose regulated protein; promoter; ss.
XX
OS Homo sapiens.
XX
PN WO200029429-A2.
XX
PD 25-MAY-2000.
XX
PF 12-NOV-1999; 99WO-JP06305.
XX
PR 13-NOV-1998; 98JP-0324227.
PR 09-JUN-1999; 99JP-0163112.
XX

```

PA (HSPR-) HSP RES INST INC.

XX Haze K, Yoshida H, Mori K, Yanagi H, Yura T;

XX WPI; 2000-387736/33.

XX New endoplasmic reticulum stress transcription factor (known as bZIP)  
PT for controlling expression of endoplasmic reticulum chaperone, useful  
PT for treating cancers, arteriosclerosis, cystic fibrosis, ischemic  
PT diseases, wounds and ulcers

XX Example 1; Fig 3; 157pp; English.

XX An endoplasmic reticulum stress transcription factor (bZIP)  
CC capable of regulating transcription inducing activity exhibited by an  
CC element (ERSE) can be used in a method for controlling expression of  
CC an endoplasmic reticulum chaperone. The method comprises expressing  
CC bZIP. The method can be used for expression of a foreign protein by  
CC positively regulating expression of an endoplasmic reticulum  
CC chaperone gene. bZIP is useful for controlling the expression of  
CC endoplasmic reticulum chaperone either positively or negatively in  
CC cells and therefore is useful for treatment or prophylaxis of  
CC cancers, arteriosclerosis, cystic fibrosis, ischaemic diseases,  
CC wounds and ulcers. bZIP also maintains the correct conformation of  
CC the endoplasmic reticulum chaperone and thereby increases the  
CC expression of a foreign protein. This sequence taken from the  
CC glucose regulating protein (GRP) promoter GRP78 contains an ERSE like  
CC sequence.

XX Sequence 19 BP; 3 A; 9 C; 5 G; 2 T; 0 other;

Query Match 52.6%; Score 10; DB 21; Length 19;

Best Local Similarity 52.6%; Pred. No. 1.8e+03;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19

Db 1 CCAATCGGCGCCTCCACG 19

RESULT 12

AAA28571

ID AAA28571 standard; DNA; 19 BP.

AC AAA28571;

XX 29-AUG-2000 (first entry)

DE GRP78 promoter ERSE1-like sequence.

XX Endoplasmic reticulum; stress; ER; transcription factor;

KW transcription; regulatory element; ERSE; bZIP; chaperone; treatment;

KW prophylaxis; cancer; arteriosclerosis; ischaemia; wound healing;

KW cystic fibrosis; ulcer; gene therapy; recombinant gene; mouse;

XX gene expression; GRP; glucose regulated protein; promoter; ss.

XX Mus musculus.

XX WO200029429-A2.

XX 25-MAY-2000.

XX 12-NOV-1999; 99WO-JP06305.

XX 13-NOV-1998; 98JP-0324227.

XX 09-JUN-1999; 99JP-0163112.

XX (HSPR-) HSP RES INST INC.

XX Haze K, Yoshida H, Mori K, Yanagi H, Yura T;

XX WPI; 2000-387736/33.

XX

PT

PT

PT

PT

XX

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Sequence 19 BP; 4 A; 8 C; 5 G; 2 T; 0 other;

Query Match 52.6%; Score 10; DB 21; Length 19;

Best Local Similarity 52.6%; Pred. No. 1.8e+03;

Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19

Db 1 CCAATCGGCGCCTCCACG 19

RESULT 13

AAA28572

ID AAA28572 standard; DNA; 19 BP.

AC AAA28572;

XX 29-AUG-2000 (first entry)

DE GRP78 promoter ERSE1-like sequence.

XX Endoplasmic reticulum; stress; ER; transcription factor;

KW transcription; regulatory element; ERSE; bZIP; chaperone; treatment;

KW prophylaxis; cancer; arteriosclerosis; ischaemia; wound healing;

KW cystic fibrosis; ulcer; gene therapy; recombinant gene; rat;

XX gene expression; GRP; glucose regulated protein; promoter; ss.

XX Rattus rattus.

XX WO200029429-A2.

XX 25-MAY-2000.

XX 12-NOV-1999; 99WO-JP06305.

XX 13-NOV-1998; 98JP-0324227.

XX 09-JUN-1999; 99JP-0163112.

XX (HSPR-) HSP RES INST INC.

XX Haze K, Yoshida H, Mori K, Yanagi H, Yura T;

XX WPI; 2000-387736/33.

XX New endoplasmic reticulum stress transcription factor (known as bZIP)  
PT for controlling expression of endoplasmic reticulum chaperone, useful  
PT for treating cancers, arteriosclerosis, cystic fibrosis, ischemic  
PT diseases, wounds and ulcers

XX Example 1; Fig 3; 157pp; English.

XX An endoplasmic reticulum stress transcription factor (bZIP)  
 CC capable of regulating transcription inducing activity exhibited by an  
 CC element (ERSE) can be used in a method for controlling expression of  
 CC an endoplasmic reticulum chaperone. The method comprises expressing  
 CC bZIP. The method can be used for expression of a foreign protein by  
 CC positively regulating expression of an endoplasmic reticulum  
 CC chaperone gene. bZIP is useful for controlling the expression of  
 CC endoplasmic reticulum chaperone either positively or negatively in  
 CC cells and therefore is useful for treatment or prophylaxis of  
 CC cancers, arteriosclerosis, cystic fibrosis, ischaemic diseases,  
 CC wounds and ulcers. bZIP also maintains the correct conformation of  
 CC the endoplasmic reticulum chaperone and thereby increases the  
 CC expression of a foreign protein. This sequence taken from the  
 CC glucose regulating protein (GRP) promoter GRP78 contains an ERSE like  
 CC sequence.  
 SQ Sequence 19 BP; 4 A; 8 C; 5 G; 2 T; 0 other;  
 Query Match 52.6%; Score 10; DB 21; Length 19;  
 Best Local Similarity 52.6%; Pred. No. 1.8e+03;  
 Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 CCAATNNNNNNNNCCACG 19  
 ID AAAA28573  
 AC AAAA28573;  
 XX 29-AUG-2000 (first entry)  
 DT GRP94 promoter ERSE1-like sequence.  
 DE  
 XX Endoplasmic reticulum; stress; ER; transcription factor;  
 KW transcription; regulatory element; ERSE; bZIP; chaperone; treatment;  
 KW prophylaxis; cancer; arteriosclerosis; ischaemia; wound healing;  
 KW cystic fibrosis; ulcer; gene therapy; recombinant gene; human;  
 KW gene expression; GRP; glucose regulated protein; promoter; ss.  
 XX Homo sapiens.  
 OS  
 XX WO200029429-A2.  
 PN 25-MAY-2000.  
 XX 12-NOV-1999; 99WO-JP06305.  
 XX 13-NOV-1998; 98JP-0324227.  
 PR 09-JUN-1999; 99JP-0163112.  
 XX (HSPR-) HSP RES INST INC.  
 PA Haze K, Yoshida H, Mori K, Yanagi H, Yura T;  
 PI WPI; 2000-387736/33.  
 XX New endoplasmic reticulum stress transcription factor (known as bZIP)  
 PT for controlling expression of endoplasmic reticulum chaperone, useful  
 PT for treating cancers, arteriosclerosis, cystic fibrosis, ischemic  
 PT diseases, wounds and ulcers  
 XX Example 1; Fig 3; 157pp; English.  
 PS An endoplasmic reticulum stress transcription factor (bZIP)  
 CC capable of regulating transcription inducing activity exhibited by an  
 CC element (ERSE) can be used in a method for controlling expression of  
 CC an endoplasmic reticulum chaperone. The method comprises expressing  
 CC bZIP. The method can be used for expression of a foreign protein by  
 CC positively regulating expression of an endoplasmic reticulum  
 CC chaperone gene. bZIP is useful for controlling the expression of  
 CC endoplasmic reticulum chaperone either positively or negatively in  
 CC cells and therefore is useful for treatment or prophylaxis of  
 CC cancers, arteriosclerosis, cystic fibrosis, ischaemic diseases,  
 CC wounds and ulcers. bZIP also maintains the correct conformation of

CC positively regulating expression of an endoplasmic reticulum  
 CC chaperone gene. bZIP is useful for controlling the expression of  
 CC endoplasmic reticulum chaperone either positively or negatively in  
 CC cells and therefore is useful for treatment or prophylaxis of  
 CC cancers, arteriosclerosis, cystic fibrosis, ischaemic diseases,  
 CC wounds and ulcers. bZIP also maintains the correct conformation of  
 CC the endoplasmic reticulum chaperone and thereby increases the  
 CC expression of a foreign protein. This sequence taken from the  
 CC glucose regulating protein (GRP) promoter GRP94 contains an ERSE like  
 CC sequence.  
 SQ Sequence 19 BP; 4 A; 10 C; 4 G; 1 T; 0 other;  
 Query Match 52.6%; Score 10; DB 21; Length 19;  
 Best Local Similarity 52.6%; Pred. No. 1.8e+03;  
 Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 CCAATNNNNNNNNCCACG 19  
 ID AAAA28574  
 AC AAAA28574;  
 XX 29-AUG-2000 (first entry)  
 DT GRP94 promoter ERSE1-like sequence.  
 DE  
 XX Endoplasmic reticulum; stress; ER; transcription factor;  
 KW transcription; regulatory element; ERSE; bZIP; chaperone; treatment;  
 KW prophylaxis; cancer; arteriosclerosis; ischaemia; wound healing;  
 KW cystic fibrosis; ulcer; gene therapy; recombinant gene; chicken;  
 KW gene expression; GRP; glucose regulated protein; promoter; ss.  
 XX Gallus domesticus.  
 OS  
 XX WO200029429-A2.  
 PN 25-MAY-2000.  
 XX 12-NOV-1999; 99WO-JP06305.  
 XX 13-NOV-1998; 98JP-0324227.  
 PR 09-JUN-1999; 99JP-0163112.  
 XX (HSPR-) HSP RES INST INC.  
 PA Haze K, Yoshida H, Mori K, Yanagi H, Yura T;  
 PI WPI; 2000-387736/33.  
 XX New endoplasmic reticulum stress transcription factor (known as bZIP)  
 PT for controlling expression of endoplasmic reticulum chaperone, useful  
 PT for treating cancers, arteriosclerosis, cystic fibrosis, ischemic  
 PT diseases, wounds and ulcers  
 XX Example 1; Fig 3; 157pp; English.  
 PS An endoplasmic reticulum stress transcription factor (bZIP)  
 CC capable of regulating transcription inducing activity exhibited by an  
 CC element (ERSE) can be used in a method for controlling expression of  
 CC an endoplasmic reticulum chaperone. The method comprises expressing  
 CC bZIP. The method can be used for expression of a foreign protein by  
 CC positively regulating expression of an endoplasmic reticulum  
 CC chaperone gene. bZIP is useful for controlling the expression of  
 CC endoplasmic reticulum chaperone either positively or negatively in  
 CC cells and therefore is useful for treatment or prophylaxis of  
 CC cancers, arteriosclerosis, cystic fibrosis, ischaemic diseases,  
 CC wounds and ulcers. bZIP also maintains the correct conformation of

CC the endoplasmic reticulum chaperone and thereby increases the  
 CC expression of a foreign protein. This sequence taken from the  
 CC glucose regulating protein (GRP) promoter GRP94 contains an ERSE like  
 CC sequence.

XX  
 SQ Sequence 19 BP; 5 A; 7 C; 6 G; 1 T; 0 other;

Query Match 52.6%; Score 10; DB 21; Length 19;  
 Best Local Similarity 52.6%; Pred. No. 1.8e+03;  
 Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNCCACG 19  
 |||||  
 Db 1 CCAATGGGAGCGCACCAG 19  
 |||||

Search completed: December 4, 2003, 06:47:54  
 Job time : 257 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 06:00:59 ; Search time 74 Seconds  
(without alignments)  
113.328 Million cell updates/sec

Title: US-09-606-804-1

Perfect score: 19  
Sequence: 1 ccaatnnnnnnnnccacg 19

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_NA.\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	52.6	19	4	US-09-813-937-1
2	10	52.6	205	4	US-09-313-294A-1696
3	10	52.6	261	4	US-09-389-681-206
4	10	52.6	261	4	US-09-620-405B-206
5	10	52.6	261	4	US-09-339-338-206
6	10	52.6	261	4	US-09-433-826B-206
7	10	52.6	261	4	US-09-604-287A-206
8	10	52.6	270	4	US-09-313-294A-890
9	10	52.6	272	1	US-08-248-474-36
10	10	52.6	272	3	US-09-756-849-36
11	10	52.6	282	4	US-09-252-991A-987
12	10	52.6	285	4	US-09-252-991A-60
13	10	52.6	296	2	US-08-716-942-20
14	10	52.6	296	4	US-09-130-337A-20
15	10	52.6	300	4	US-09-313-294A-7348
16	10	52.6	360	6	5196523-1
17	10	52.6	387	3	US-08-804-372A-23
18	10	52.6	392	4	US-09-404-879A-372
19	10	52.6	399	6	5196523-23
20	10	52.6	421	6	5196523-22
21	10	52.6	424	1	US-08-609-657-15
22	10	52.6	432	4	US-09-252-991A-9959
23	10	52.6	457	4	US-09-702-705-861
24	10	52.6	457	4	US-09-736-457-861
25	10	52.6	467	3	US-09-257-584-8
26	10	52.6	511	4	US-09-702-705-1138
27	10	52.6	511	4	US-09-736-457-1138

c	28	10	52.6	522	4	US-09-663-600A-149	Sequence 149, App
	29	10	52.6	531	4	US-09-222-575-123	Sequence 123, App
	30	10	52.6	531	4	US-09-389-681-123	Sequence 123, App
	31	10	52.6	531	4	US-09-620-405B-123	Sequence 123, App
	32	10	52.6	531	4	US-09-339-338-123	Sequence 123, App
	33	10	52.6	531	4	US-09-433-826B-123	Sequence 123, App
	34	10	52.6	531	4	US-09-604-287A-123	Sequence 123, App
	35	10	52.6	584	4	US-09-663-600A-55	Sequence 55, Appl
	36	10	52.6	641	4	US-09-573-906-7	Sequence 7, Appl
	37	10	52.6	708	4	US-09-252-991A-3364	Sequence 3364, Ap
	38	10	52.6	754	3	US-09-020-956-20	Sequence 20, Appl
	39	10	52.6	754	3	US-09-030-607-20	Sequence 20, Appl
	40	10	52.6	754	4	US-09-439-313-20	Sequence 20, Appl
	41	10	52.6	754	4	US-09-352-616A-20	Sequence 20, Appl
	42	10	52.6	754	4	US-09-232-149A-20	Sequence 20, Appl
	43	10	52.6	770	4	US-09-573-906-6	Sequence 6, Appl
	44	10	52.6	843	4	US-09-171-209-38	Sequence 38, Appl
	45	10	52.6	849	3	US-08-998-416-552	Sequence 552, App

ALIGNMENTS

RESULT 1  
US-09-813-937-1  
; Sequence 1, Application US/09813937  
; Patent No. 6506600  
; GENERAL INFORMATION:  
; APPLICANT: HERMONAT, Paul L.  
; APPLICANT: MANE, Michael  
; APPLICANT: LIU, Yong  
; TITLE OF INVENTION: SECRETING PRODUCTS FROM SKIN BY ADENO-ASSOCIATED VIRUS (AAV) GEN  
; TITLE OF INVENTION: TRANSFER  
; FILE REFERENCE: 023533/0116  
; CURRENT APPLICATION NUMBER: US/09/813.937  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: US 60/191092  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 1  
; LENGTH: 19  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: glucose response element  
; NAME/KEY: misc feature  
; LOCATION: (6)-(14)  
; OTHER INFORMATION: "N" at positions 6 - 14 can be A, C, G or T  
US-09-813-937-1

Query Match 52.6%; Score 10; DB 4; Length 19;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CCAATNNNNNNNNCCACG 19  
Db 1 CCAATNNNNNNNNCCACG 19

RESULT 2  
US-09-313-294A-1696/c  
; Sequence 1696, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalugudi, Raghunath V.  
; APPLICANT: Ito, Laura I.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313.294A  
; CURRENT FILING DATE: 1999-05-14

```
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 1696
; LENGTH: 205
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700551354H1
US-09-313-294A-1696

Query Match      52.6%; Score 10; DB 4; Length 205;
Best Local Similarity 52.6%; Pred. No. 5.2e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 122 CCAATCAGACAGCCACG 104

RESULT 3
US-09-389-681-206
; Sequence 206, Application US/09389681A
; Patent No. 6518237
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C3
; CURRENT APPLICATION NUMBER: US/09/389,681A
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-389-681-206

Query Match      52.6%; Score 10; DB 4; Length 261;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 140 CCAATTCCTCATCTCCACG 158

RESULT 4
US-09-620-405B-206
; Sequence 206, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-620-405B-206

Query Match      52.6%; Score 10; DB 4; Length 261;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 140 CCAATTCCTCATCTCCACG 158

RESULT 5
US-09-339-338-206
; Sequence 206, Application US/09339338A
; Patent No. 6573368
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C2
; CURRENT APPLICATION NUMBER: US/09/339,338A
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-339-338-206

Query Match      52.6%; Score 10; DB 4; Length 261;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 140 CCAATTCCTCATCTCCACG 158

RESULT 6
US-09-433-826B-206
; Sequence 206, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-433-826B-206

Query Match      52.6%; Score 10; DB 4; Length 261;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNNNCCACG 19
Db 140 CCAATTCCTCATCTCCACG 158
```



```
RESULT 7
US-09-604-287A-206
; Sequence 206, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-604-287A-206

Query Match 52.6%; Score 10; DB 4; Length 261;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19
Db 140 CCAATTCCTCATCTCCACG 158

RESULT 8
US-09-313-294A-890/c
; Sequence 890, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 890
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700550018H1
; NAME/KEY: unsure
; LOCATION: 213-215, 218-221, 232-233, 236-237, 239-240, 242, 249, 251, 253, 261,
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-890

Query Match 52.6%; Score 10; DB 4; Length 270;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19
Db 121 CCAATCAGACAGACCCACG 103

RESULT 9
US-08-248-474-36
; Sequence 36, Application US/08248474
; Patent No. 5612471
```

```
; GENERAL INFORMATION:
; APPLICANT: MCK, BIRD, David
; APPLICANT: WILSON, Mark A.
; TITLE OF INVENTION: NEMATODE-INDUCE GENES IN TOMATO
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,474
; FILING DATE: 25-MAY-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 2307E-535
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Lycopersicon esculentum cv 'Rutgers Large
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..272
; OTHER INFORMATION: /standard_name= "DB# 155"
US-08-248-474-36

Query Match 52.6%; Score 10; DB 1; Length 272;
Best Local Similarity 52.6%; Pred. No. 5.4e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCAATNNNNNNCCACG 19
Db 97 CCAATATTATTACCACG 115

RESULT 10
US-08-756-849-36
; Sequence 36, Application US/08756849
; Patent No. 6093810
; GENERAL INFORMATION:
; APPLICANT: Bird, David Mck.
; APPLICANT: Wilson, Mark A.
; TITLE OF INVENTION: Nematode-Induced Genes in Tomato
; NUMBER OF SEQUENCES: 129
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/756,849  
; FILING DATE: 26-NOV-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/248,474  
; FILING DATE: 25-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 023070-053510US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 272 base pairs  
; TYPE: nucleic acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: Lycopersicon esculentum cv 'Rutgers Large Red'  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1..272  
; OTHER INFORMATION: /standard\_name= "DB# 155"  
US-08-756-849-36

Query Match 52.6%; Score 10; DB 3; Length 272;  
Best Local Similarity 52.6%; Pred. No. 5.4e+02;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
Db 97 CCAATATTATTTTACCACG 115

RESULT 11  
US-09-252-991A-987  
; Sequence 987, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 987  
; LENGTH: 282  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-987

Query Match 52.6%; Score 10; DB 4; Length 282;  
Best Local Similarity 52.6%; Pred. No. 5.4e+02;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
Db 49 CCAATGCTGATCCCCACG 67

RESULT 12  
US-09-252-991A-60  
; Sequence 60, Application US/09252991A

; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 60  
; LENGTH: 285  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-60

Query Match 52.6%; Score 10; DB 4; Length 285;  
Best Local Similarity 52.6%; Pred. No. 5.5e+02;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
Db 224 CCAATGACCGATACCACG 242

RESULT 13  
US-08-716-942-20/C  
; Sequence 20, Application US/08716942  
; Patent No. 5849491  
; GENERAL INFORMATION:  
; APPLICANT: Terragen Diversity Inc.  
; APPLICANT: Radowski, Christopher C. A.  
; APPLICANT: Seow, Kah Tong  
; APPLICANT: Warren, R. Antony J.  
; APPLICANT: Yap, Wai Ho  
; TITLE OF INVENTION: METHOD FOR ISOLATING XYLANASE GENE  
; TITLE OF INVENTION: SEQUENCES FROM SOIL DNA, COMPOSITIONS USEFUL IN SUCH METHOD AT  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Oppedahl & Larson  
; STREET: 1992 Commerce Street, Suite 309  
; CITY: Yorktown Heights  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10598-4412  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS 5.0  
; SOFTWARE: Word Perfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/716,942  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/004,157  
; FILING DATE: 20-Sep-95  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Marina T. Larson  
; REGISTRATION NUMBER: 32,038  
; REFERENCE/DOCKET NUMBER: TERR.P-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (914) 245-3252  
; TELEFAX: (914) 962-4330  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 296  
; TYPE: nucleic acid

```

; STRANDEDNESS: DOUBLE
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM:
; FEATURE:
; NAME/KEY: fragment of xylanase gene from degenerate primer
; NAME/KEY: amplification of soil DNA
US-08-716-942-20

Query Match          52.6%; Score 10; DB 2; Length 296;
Best Local Similarity 52.6%; Pred. No. 5.5e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
Db 217 CCAATGGTGTGGCCACG 199

RESULT 14
US-09-130-337A-20/c
; Sequence 20, Application US/09130337A
; Patent No. 6441148
; GENERAL INFORMATION:
; APPLICANT: Radomski, CCA
; APPLICANT: Seow, KT
; APPLICANT: Warren, RAJ
; APPLICANT: Yad, WH
; TITLE OF INVENTION: METHOD FOR AMPLIFICATION OF XYLANASE GENE FRAGMENTS
; TITLE OF INVENTION: FROM SOIL DNA, COMPOSITIONS USEFUL IN SUCH METHOD AND COMPOSITION
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 9993-004
; CURRENT APPLICATION NUMBER: US/09/130,337A
; CURRENT FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 08/716,942
; PRIOR FILING DATE: 1996-09-20
; PRIOR APPLICATION NUMBER: 60/004,157
; PRIOR FILING DATE: 1995-09-22
; NUMBER OF SEQ ID NOS: 25
; SEQ ID NO 20
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Unknown organism
; FEATURE:
; OTHER INFORMATION: Description of unknown organism: soil microbe
US-09-130-337A-20

Query Match          52.6%; Score 10; DB 4; Length 296;
Best Local Similarity 52.6%; Pred. No. 5.5e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
Db 217 CCAATGGTGTGGCCACG 199

RESULT 15
US-09-313-294A-7348/c
; Sequence 7348, Application US/0913294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program

```

```

; SEQ ID NO 7348
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700381587H1
; NAME/KEY: unsure
; LOCATION: 18, 228, 295
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-7348

Query Match          52.6%; Score 10; DB 4; Length 300;
Best Local Similarity 52.6%; Pred. No. 5.5e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
Db 292 CCAATCGTACCCTCCACG 274

Search completed: December 4, 2003, 07:52:25
Job time : 77 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 07:13:09 ; Search time 259 Seconds  
(without alignments)  
243.816 Million cell updates/sec

Title: US-09-606-804-1

Perfect score: 19  
Sequence: 1 ccaatnnnnnnnnccacg 19

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 166179599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	52.6	19	14	US-10-340-759-1
2	10	52.6	24	11	US-09-940-185-207
3	10	52.6	25	9	US-09-402-100-35
4	10	52.6	25	11	US-09-940-185-4188
5	10	52.6	100	10	US-09-969-373-1152
6	10	52.6	123	10	US-09-867-701-3962
7	10	52.6	136	12	US-09-922-449B-18
8	10	52.6	136	12	US-09-922-449B-19
9	10	52.6	142	12	US-09-922-449B-8
10	10	52.6	143	12	US-09-922-449B-24
11	10	52.6	149	12	US-09-922-449B-25
12	10	52.6	150	12	US-09-922-449B-10
13	10	52.6	167	14	US-10-001-883-5
14	10	52.6	210	10	US-09-974-300-8096
15	10	52.6	214	9	US-09-778-320-59
16	10	52.6	214	9	US-09-910-689-59

C	17	10	52.6	214	13	US-10-010-742-59	Sequence 59, Appl
	18	10	52.6	219	10	US-09-783-590-10136	Sequence 10136, A
	19	10	52.6	225	14	US-10-060-036-3284	Sequence 3284, Ap
	20	10	52.6	237	9	US-09-777-564-1013	Sequence 1013, Ap
	21	10	52.6	237	14	US-10-015-219-1013	Sequence 1013, Ap
C	22	10	52.6	238	10	US-09-960-352-13155	Sequence 13155, A
C	23	10	52.6	247	10	US-09-878-574-6181	Sequence 6181, Ap
	24	10	52.6	261	9	US-09-604-287A-206	Sequence 206, App
	25	10	52.6	261	10	US-09-339-338-206	Sequence 206, App
	26	10	52.6	261	11	US-09-551-621-206	Sequence 206, App
	27	10	52.6	261	12	US-10-124-805-206	Sequence 206, App
	28	10	52.6	261	13	US-10-007-805-206	Sequence 206, App
	29	10	52.6	261	14	US-10-076-622-206	Sequence 206, App
	30	10	52.6	262	12	US-10-099-926-1836	Sequence 1836, Ap
	31	10	52.6	262	13	US-10-033-528-1836	Sequence 1836, Ap
	32	10	52.6	267	9	US-09-294-093B-1896	Sequence 1896, Ap
	33	10	52.6	268	9	US-09-923-876-5857	Sequence 5857, Ap
C	34	10	52.6	271	10	US-09-783-590-11139	Sequence 11139, A
C	35	10	52.6	272	11	US-09-835-976B-57	Sequence 57, Appl
C	36	10	52.6	274	10	US-09-867-701-5278	Sequence 5278, Ap
	37	10	52.6	280	10	US-09-878-574-246	Sequence 246, App
	38	10	52.6	283	9	US-09-294-093B-3259	Sequence 3259, Ap
	39	10	52.6	293	9	US-09-294-093B-6179	Sequence 6179, Ap
C	40	10	52.6	301	11	US-09-918-995-24371	Sequence 24371, A
	41	10	52.6	306	10	US-09-974-300-8100	Sequence 8100, Ap
C	42	10	52.6	339	10	US-09-878-574-1186	Sequence 1186, Ap
	43	10	52.6	348	11	US-09-899-495-18	Sequence 18, Appl
C	44	10	52.6	351	10	US-09-938-842A-1451	Sequence 1451, Ap
	45	10	52.6	351	10	US-09-796-692-6226	Sequence 6226, Ap

## ALIGNMENTS

## RESULT 1

US-10-340-759-1  
; Sequence 1, Application US/10340759  
; Publication No. US20030104575A1  
; GENERAL INFORMATION:  
; APPLICANT: HERMONAT, Paul L.  
; APPLICANT: MANE, Michael  
; APPLICANT: LIU, Yong  
; TITLE OF INVENTION: SECRETING PRODUCTS FROM SKIN BY ADENO-ASSOCIATED VIRUS (AAV) GEN  
; FILE REFERENCE: 023533/0116  
; CURRENT APPLICATION NUMBER: US/10/340,759  
; CURRENT FILING DATE: 2003-01-13  
; PRIOR APPLICATION NUMBER: US/09/813,937A  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: US 60/191092  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 1  
; LENGTH: 19  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: glucose response element  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (6)-(14)  
; OTHER INFORMATION: "N" at positions 6 - 14 can be A, C, G or T  
US-10-340-759-1

Query Match 52.6%; Score 10; DB 14; Length 19;  
Best Local Similarity 100.0%; Pred. No. 4.1e+03;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CCAATNNNNNNNNCCACG 19  
Db 1 CCAATNNNNNNNNNNCCACG 19

```

RESULT 2
US-09-940-185-207/c
; Sequence 207, Application US/09940185
; Publication No. US20030096239A1
; GENERAL INFORMATION:
; APPLICANT: Gunderson, Kevin
; TITLE OF INVENTION: Probes and Decoder Oligonucleotides
; FILE REFERENCE: A-69605-1
; CURRENT APPLICATION NUMBER: US/09/940,185
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/227,948
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/228,854
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 4768
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 207
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Computer Generated Probe Sequence.
US-09-940-185-207

Query Match          52.6%; Score 10; DB 11; Length 24;
Best Local Similarity 52.6%; Pred. No. 4.2e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 23 CCAATATTACGTGACCACG 5

RESULT 3
US-09-402-100-35/c
; Sequence 35, Application US/09402100
; Patent No. US20010019834A1
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co, LTD
; APPLICANT: Kim, Byung-O
; APPLICANT: Shin, Sung-Seup
; APPLICANT: Yu, Young-Hyo
; APPLICANT: Park, Myung-Hwan
; APPLICANT: Choi, Deok-Joon
; APPLICANT: Jung, Hyung-Jin
; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Helicobacter
; FILE REFERENCE: 0136/OG140
; CURRENT APPLICATION NUMBER: US/09/402,100
; CURRENT FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: KR 97-11950
; EARLIER FILING DATE: 1997-03-31
; EARLIER APPLICATION NUMBER: KR 97-11951
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: ( )..(?)
; OTHER INFORMATION: Oligonucleotide
US-09-402-100-35

Query Match          52.6%; Score 10; DB 9; Length 25;
Best Local Similarity 52.6%; Pred. No. 4.2e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 23 CCAATATTACGTGACCACG 5

RESULT 4
US-09-940-185-4188/c
; Sequence 4188, Application US/09940185
; Publication No. US20030096239A1
; GENERAL INFORMATION:
; APPLICANT: Gunderson, Kevin
; TITLE OF INVENTION: Probes and Decoder Oligonucleotides
; FILE REFERENCE: A-69605-1
; CURRENT APPLICATION NUMBER: US/09/940,185
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/227,948
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/228,854
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 4768
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4188
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Computer Generated Probe Sequence.
US-09-940-185-4188

Query Match          52.6%; Score 10; DB 11; Length 25;
Best Local Similarity 52.6%; Pred. No. 4.2e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 24 CCAATATTACGTGACCACG 6

RESULT 5
US-09-969-373-1152/c
; Sequence 1152, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
; APPLICANT: Effertz, Roger J.
; APPLICANT: Hauge, Brian M.
; TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
; FILE REFERENCE: 38-10(52679)A
; CURRENT APPLICATION NUMBER: US/09/969,373
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 09/754,853
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 09/760,427
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 09/855,768
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 1152
; LENGTH: 100
; TYPE: DNA
; ORGANISM: Glycine max
US-09-969-373-1152

Query Match          52.6%; Score 10; DB 10; Length 100;
Best Local Similarity 52.6%; Pred. No. 4.9e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19
    |||||
Db 97 CCAATCATTTCCATCCACG 79

RESULT 6
US-09-867-701-3962
; Sequence 3962, Application US/09867701
```

```
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3962
; LENGTH: 123
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-3962

Query Match      52.6%; Score 10; DB 10; Length 123;
Best Local Similarity 52.6%; Pred. No. 5e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 CCAATNNNNNNNNCCACG 19
      |||||
Db      4 CCAATCTGTCGGACCACG 22

RESULT 7
US-09-922-449B-18
; Sequence 18, Application US/09922449B
; Publication No. US20030148278A1
; GENERAL INFORMATION:
; APPLICANT: BioInside Gesellschaft fur Biodiagnostik, Auftragsforschung und Consulting
; APPLICANT: mbH
; TITLE OF INVENTION: Test kit and method for quantitatively detecting genetically modified
; TITLE OF INVENTION: in foodstuff by means of fluorescent-coupled PCR
; FILE REFERENCE: 101215-68
; CURRENT APPLICATION NUMBER: US/09/922,449B
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: PCT/EP00/009835
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: DE 199 06 169.6
; PRIOR FILING DATE: 1999-02-08
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 136
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: target IAC DNA for the Bt-176
US-09-922-449B-18

Query Match      52.6%; Score 10; DB 12; Length 136;
Best Local Similarity 52.6%; Pred. No. 5.1e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 CCAATNNNNNNNNCCACG 19
      |||||
Db      27 CCAATTTTCGCCTCCACG 45

RESULT 8
US-09-922-449B-19
; Sequence 19, Application US/09922449B
; Publication No. US20030148278A1
; GENERAL INFORMATION:
; APPLICANT: BioInside Gesellschaft fur Biodiagnostik, Auftragsforschung und Consulting
; APPLICANT: mbH
; TITLE OF INVENTION: Test kit and method for quantitatively detecting genetically modified
; TITLE OF INVENTION: in foodstuff by means of fluorescent-coupled PCR
; FILE REFERENCE: 101215-68
```

```
; CURRENT APPLICATION NUMBER: US/09/922,449B
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: PCT/EP00/009835
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: DE 199 06 169.6
; PRIOR FILING DATE: 1999-02-08
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 136
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: reference IAC DNA for the Bt-176
US-09-922-449B-19

Query Match      52.6%; Score 10; DB 12; Length 136;
Best Local Similarity 52.6%; Pred. No. 5.1e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 CCAATNNNNNNNNCCACG 19
      |||||
Db      27 CCAATTTTCGCCTCCACG 45

RESULT 9
US-09-922-449B-8
; Sequence 8, Application US/09922449B
; Publication No. US20030148278A1
; GENERAL INFORMATION:
; APPLICANT: BioInside Gesellschaft fur Biodiagnostik, Auftragsforschung und Consulting
; APPLICANT: mbH
; TITLE OF INVENTION: Test kit and method for quantitatively detecting genetically modified
; TITLE OF INVENTION: in foodstuff by means of fluorescent-coupled PCR
; FILE REFERENCE: 101215-68
; CURRENT APPLICATION NUMBER: US/09/922,449B
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: PCT/EP00/009835
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: DE 199 06 169.6
; PRIOR FILING DATE: 1999-02-08
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 142
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: target IAC DNA for the RRS 9
US-09-922-449B-8

Query Match      52.6%; Score 10; DB 12; Length 142;
Best Local Similarity 52.6%; Pred. No. 5.1e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 CCAATNNNNNNNNCCACG 19
      |||||
Db      31 CCAATTTTCGCCTCCACG 49

RESULT 10
US-09-922-449B-24
; Sequence 24, Application US/09922449B
; Publication No. US20030148278A1
; GENERAL INFORMATION:
; APPLICANT: BioInside Gesellschaft fur Biodiagnostik, Auftragsforschung und Consulting
; APPLICANT: mbH
; TITLE OF INVENTION: Test kit and method for quantitatively detecting genetically modified
; TITLE OF INVENTION: in foodstuff by means of fluorescent-coupled PCR
; FILE REFERENCE: 101215-68
; CURRENT APPLICATION NUMBER: US/09/922,449B
; CURRENT FILING DATE: 2001-08-03
```

; PRIOR APPLICATION NUMBER: PCT/EP00/009835  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: DE 199 06 169.6  
; PRIOR FILING DATE: 1999-02-08  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 24  
; LENGTH: 143  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: target IAC DNA for the RRS ge  
US-09-922-449B-24

Query Match 52.6%; Score 10; DB 12; Length 143;  
Best Local Similarity 52.6%; Pred. No. 5.1e+03;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 31 CCAATTTTCGCCTCCACG 49

## RESULT 11

US-09-922-449B-25  
; Sequence 25, Application US/09922449B  
; Publication No. US20030148278A1  
; GENERAL INFORMATION:  
; APPLICANT: Bioinside Gesellschaft fur Bodiagnostik, Auftragsforschung und Consulting  
; APPLICANT: mbH  
; TITLE OF INVENTION: Test kit and method for quantitatively detecting genetically modified  
; FILE OF INVENTION: in foodstuff by means of fluorescent-coupled PCR  
; FILE REFERENCE: 101215-68  
; CURRENT APPLICATION NUMBER: US/09/922,449B  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: PCT/EP00/009835  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: DE 199 06 169.6  
; PRIOR FILING DATE: 1999-02-08  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 25  
; LENGTH: 149  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: reference IAC DNA for the RRS  
US-09-922-449B-25

Query Match 52.6%; Score 10; DB 12; Length 149;  
Best Local Similarity 52.6%; Pred. No. 5.1e+03;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 31 CCAATTTTCGCCTCCACG 49

## RESULT 12

US-09-922-449B-10  
; Sequence 10, Application US/09922449B  
; Publication No. US20030148278A1  
; GENERAL INFORMATION:  
; APPLICANT: Bioinside Gesellschaft fur Bodiagnostik, Auftragsforschung und Consulting  
; APPLICANT: mbH  
; TITLE OF INVENTION: Test kit and method for quantitatively detecting genetically modified  
; FILE OF INVENTION: in foodstuff by means of fluorescent-coupled PCR  
; FILE REFERENCE: 101215-68  
; CURRENT APPLICATION NUMBER: US/09/922,449B  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: PCT/EP00/009835  
; PRIOR FILING DATE: 2000-02-07

; PRIOR APPLICATION NUMBER: DE 199 06 169.6  
; PRIOR FILING DATE: 1999-02-08  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 150  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: reference IAC DNA for the RRS  
; OTHER INFORMATION: gen  
US-09-922-449B-10

Query Match 52.6%; Score 10; DB 12; Length 150;  
Best Local Similarity 52.6%; Pred. No. 5.1e+03;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 32 CCAATTTTCGCCTCCACG 50

## RESULT 13

US-10-001-883-5/c  
; Sequence 5, Application US/10001883  
; Publication No. US20030022188A1  
; GENERAL INFORMATION:  
; APPLICANT: Macina, Roberto  
; APPLICANT: Recipon, Hervé  
; APPLICANT: Pluta, Jason  
; APPLICANT: Ghosh, Malavika  
; APPLICANT: Sun, Yongming  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Pri  
; FILE REFERENCE: DEX-0271  
; CURRENT APPLICATION NUMBER: US/10/001,883  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 60/252,059  
; PRIOR FILING DATE: 2000-11-20  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 167  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-001-883-5

Query Match 52.6%; Score 10; DB 14; Length 167;  
Best Local Similarity 52.6%; Pred. No. 5.2e+03;  
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCAATNNNNNNNNCCACG 19  
|||||  
Db 160 CCAATTTTGAATCACCACG 142

## RESULT 14

US-09-974-300-8096/c  
; Sequence 8096, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; APPLICANT: Clausen, Ib Groth  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; FILE OF INVENTION: Expression  
; FILE REFERENCE: 10085.500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; CURRENT FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/279,526  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 8481

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8096
; LENGTH: 210
; TYPE: DNA
; ORGANISM: Bacillus clausii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(210)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-8096

Query Match      52.6%; Score 10; DB 10; Length 210;
Best Local Similarity 52.6%; Pred.No. 5.3e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy      1 CCAATNNNNNNNCCACG 19
      |||||
Db      152 CCAATTTCGTCTGCCACG 134

RESULT 15
US-09-778-320-59/c
; Sequence 59, Application US/09778320
; Patent No. US20010034052A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Day, Craig H.
; APPLICANT: Jiang, Yugu
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Wang, TongTong
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.491CS
; CURRENT APPLICATION NUMBER: US/09/778,320
; CURRENT FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 214
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(214)
; OTHER INFORMATION: n = A,T,C or G
US-09-778-320-59

Query Match      52.6%; Score 10; DB 9; Length 214;
Best Local Similarity 52.6%; Pred.No. 5.3e+03;
Matches 10; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy      1 CCAATNNNNNNNCCACG 19
      |||||
Db      88 CCAATTCTTCATCTCCACG 70
```

Search completed: December 4, 2003, 08:51:36  
Job time : 263 secs